THE UNIVERSITY
of EDINBURGH

Breast Cancer Heterogeneity within the Neoadjuvant Window

Dominic Pearce

Doctor of Philosophy
The University of Edinburgh
2018

Abstract

Breast cancer affects ~55,000 women in the UK annually. Key to the reliable management of the disease is robust characterisation of an individual's tumour, to best determine the most appropriate treatment and likely prognosis. High-throughput measurement of gene expression has driven recent developments in improved clinical patient stratification but difficulties in accuracy and sensitivity persist. Key to producing — and validating — predictive signatures of disease response and progression lies in utilising patient-derived material, to maximise a study's relevance to clinical practice. Traditionally, measurements taken at diagnosis or surgery have been correlated with long-term outcomes, such as evidence of recurrence or metastasis. These post-operative studies often take place over the course of decades and are understandably restricted in terms of speed, efficiency and cost. In contrast, the *pre-operative* neoadjuvant setting allows expedited, short course studies that maximise the information available from a tumour whilst it is still *in situ*, allowing multiple characterisations of the tumour to be performed over time. The *in situ* nature of neoadjuvant study has allowed phenotypic and molecular characterisation to be performed in concert, allowing, for example, treatment response to be explained in terms of variation in gene expression. Supporting these neoadjuvant studies are routine biobanking operations, facilitating repeated tumour sampling and physiological characterisations. A happy side-effect of this indiscriminate sampling strategy is the generation of cohorts that do not fit the traditional neoadjuvant model and offer the potential to query alternative hypotheses. This thesis makes use of two of these cohorts to investigate each of (i) intra-tumour heterogeneity under conditions of no-treatment and (ii) long-term latent treatment resistance. Key results include demonstrating a biopsy method-specific effect on gene expression and revealing a role of epigenetic modification in endocrine treatment resistance, respectively. In addition, a more generally applicable methodology to illustrate and quantify the association of a continuous variable — such as gene expression — with outcome is described. Taken together, these parallel threads depict the emerging utility of the neoadjuvant setting in portraying difficult to model clinically relevant aspects of cancer treatment and response. The results will likely prove to guide clinical best practice as well as inform future studies, with the novel datasets generated allowing comparison, validation and further analysis.

Lay abstract

Just as the weather can be forecast using patterns of temperature, humidity and pressure, cancer can be diagnosed by measuring patterns within our bodies. The hope is to be able to tell if a person's cancer will respond or not respond to treatment. Just like with a weather forecast, this allows us to make decisions of how to act. We can both choose whether to treat when necessary (or wear a raincoat), as well as when not to treat (so that we aren't caught wearing three jumpers on the hottest day of the year). Here we are forecasting the likely course of events but, as anyone caught in a sudden shower will know, weather forecasts, like cancer forecasts, are often not always reliable. More recently, new technologies have enabled these measurements to be made on smaller and smaller parts of us, even to the point of looking at what is happening within a single cell. It is hoped that these finer measurements will allow for more accurate forecasts but, just as we would have to wait a week or so to check whether a weather forecast is correct, we often have to wait decades to check whether a cancer forecast is correct or not. This is like waiting to check if it rained or not by seeing if a lake had formed 100 years in the future. The reason for this wait is because the tumour is removed at the time the forecast is made and there is nothing left to check whether the treatment is working or not. Not only is this delay far too long a time, the cost of running such a study is large. For cancer, forecasting can be sped up by taking measurements earlier. Here we can look at how these measurements compare to how the tumour responds whilst it is still in place. We can then find the similarities between people whose tumours respond, and those between people whose tumour did not respond. Knowing these similarities we can test a new cancer and predict whether they are more similar to the responders or the non-responders. Unfortunately, cancer treatment is more complicated than simply whether someone will respond or not, and patients often respond to treatment at first before becoming resistant. In this thesis we will use early tumour measurements to help guide how cancer forecasts should best be made, asking - (i) can resistant tumours be forecast and (iii) what happens when we forecast tumours that are untreated. Together, these threads show the benefit of using early tumour measurements and how they may be used by doctors and health professionals.

Declaration

I, Dominic Pearce, confirm that the work presented in this thesis is my own and has not been submitted elsewhere for any other degree or professional qualification. Where information has been derived from other sources or in collaboration with colleagues, I confirm that this has been indicated.

Old bard I like you more
    now that I know you're
    no Saturday Evening Post philosopher
Nay but such who plagarizes God
                            - Gregory Corso

Acknowledgements

Beyond the obvious of my parents, brothers, friends and co-workers, this thesis is indebted to three persons whose influences have shaped it the most. Firstly, my supervisor Andy, it goes without saying that I couldn't have done it, and wouldn't have known how to do it, without your guidance and confidence to allow me to be thrown head first into a subject I had little (or no) experience with. Your calm support and inclination to allow me to figure problems out by myself has fed a self-sufficiency worth far more than the contents of this thesis. Secondly, Félix, who's timing could not have been worse, and whose squeals, cries and fractious movements have mirrored my own during these final few months. I'm glad that you made it before it was all over and that your fingerprints — currently figurative, soon to be literal — are found throughout. Most importantly, my wife Kat, who has helped me cope — and coped with me — for each and every success and failure, who has kept my life whole and full of happiness and endeavour, we made it!

This thesis would have been impossible without the work performed in collaboration with my own, not least the tireless efforts of the Edinburgh Breast Clinic, in organising, collecting and processing the wealth of data that has fed my own work. This is equally true for the lab of Professor Robert Clarke, who, beyond providing Chapter 4's validation dataset, took care of me beyond what I could have expected during my time with them at Georgetown University. Finally, a special thanks to the hundreds of women who agreed for their own persons, discomfort and effort to be put towards research in hope of improving breast cancer treatment and its management for others.

Thesis structure & contributions

This thesis is composed of six sections: an introduction, general methods, three results chapters and a general discussion. In addition to the overarching introduction and discussion, each results chapter includes two small introductory and concluding sections to add study-specific context and perspectives. Each results chapter is is composed of previously published work[1–3] that has been reworked for the purposes of this thesis, including additional figures and analysis. I present here a brief outline (with thanks) of the contributions made in collaboration during these investigations.

Tissue collection : J. Michael Dixon & Lorna Renshaw
Pathology : Jeremy Thomas & John Bartlett
Contributions to lab work : Cigdem Selli, Vicky Sabine & Anu Fernando
Contributions to data analysis : Andy Sims, Cidgem Selli, Gil Tomàs, Arran Turnbull, Laura Arthur, Ian Simpson & J. Michael Dixon
Validation data generation (Georgetown University, Washington D.C.) : The laboratory of Professor Robert Clarke

Funding for this project was provided by Cancer Research UK.

Contents

1 | Introduction

Cancer is a dysregulation of otherwise normal function necessary to develop and preserve life.[i] It corrupts the same biochemical complexities that grant organisms greater diversity, survival and fitness,[ii] complexity that ultimately provides opportunity for perturbation. In simpler terms, the more complex the system becomes the more challenging it is to keep the scales of growth, survival, and death balanced appropriately[6]. This is captured in the widely cited cancer hallmarks[7], with dysregulation of growth signalling, apoptosis, replication, angiogenesis, and tissue invasion all representing otherwise normal cellular functions repurposed within an aberrant context. At the heart of this *normality* of tumourigenesis is a treatment paradox — how are we able to treat a tumour in lieu of ourselves.

1.1 Cancer therapy

Treatment for solid tumours can be approached either locally — by surgery and radiotherapy — as well as systemically, including chemo-, immuno- and hormone therapies. Systemic chemotherapy has revolved around targeting inherent susceptibilities common to a majority of cancers,[iii] taking advantage of characteristic high rates of proliferation and impaired DNA damage repair pathways[8]. It is therefore at the point of DNA replication and cellular division that tumour cells become most sensitive to the DNA damaging and anti-mitotic agents that compose the bulk of chemotherapeutics. However, whilst these agents are extremely effective at eliminating tumour cells, their off-target effects on other susceptible tissues are severely debilitating[9]. This limits their use, with sub-optimal levels defined by a 'maximally tolerated dose', necessary to balance effect and side-effect. We can consider avoiding treatment-induced side-effects and achieving this balance by two complimentary approaches: in the development of tumour-specific targeted therapies; and in predicting response to treatment whilst avoiding over treatment. For the purposes of this

[i] "We have only seen our monster more clearly and described his scales and fangs in new ways – ways that reveal a cancer cell to be, like Grendel, a distorted version of our normal selves".[4]

[ii] In fact, current understanding suggests that cancer has paralleled multicellularity since the earliest metazoa.[5]

[iii] Notable difficulties in administering systemic treatment include brain tumours protected by the blood-brain barrier.

thesis we will consider the latter, emphasising the idea that targeted therapies are of little use if we don't know where to aim them.

1.2 Breast cancer

Over the course of the 21st century, breast cancer has stood out as emblematic in the development of novel bioinformatics approaches towards prognosis and treatment prediction for a number of reasons[10]. Not least amongst these are its high incidence, relative ease in procuring biopsy material and popular research support from both charitable and governmental funding bodies[6]. Together, these factors have enabled and encouraged extensive sample collection, allowing increasingly bold studies to be realised[11]. In this way breast cancer has pioneered advancements in high-throughput molecular methodology and fuelled its own molecular characterisation.

As a disease, breast cancer presents as a solid tumour in ~55,000 UK women per year. Despite decreasing mortality, incidence continues to rise, leaving breast cancer responsible for ~11,500 deaths annually, representing ~9% of cancer-related deaths in the UK[12]. Improved screening is understood to partially explain this paradox of increasing incidence and decreasing mortality, and is suggestive of two concepts — that earlier characterisation of a tumour allows better treatment, and that not all cancers detected *in situ* progress to become invasive and require treatment.[i] Disentangling which tumours are — or will become — invasive and/or recurrent/metastatic, and which will remain benign is critical to the full realisation of diagnostic screening efforts. Here again is highlighted the concept of treating those tumours that require it, whilst avoiding over treating patients who do not. Decisions to treat or not treat, or how best to treat, have proven challenging for many cancers, with stratification schemes consistently falling short in accuracy. Even marginal shortfalls in the context of a highly prevalent disease can translate to thousands of women being incorrectly treated per year, and have therefore motivated continued attempts to improve prognostic and predictive accuracy.[ii] These attempts have

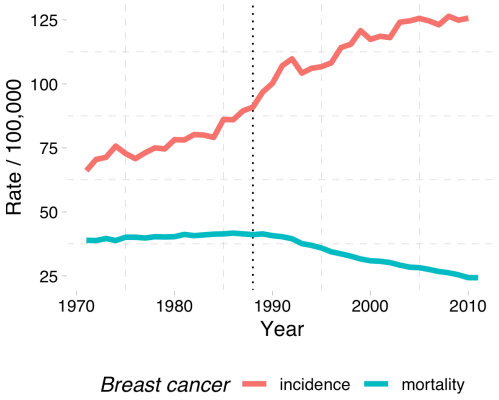

Figure 1.2.1: UK breast cancer incidence and mortality rates per 100,000 women. Source: Office for National Statistics.

[i] H Gilbert Welch adds that mortality and incidence may be even more closely entangled, "the more overdiagnosis the test causes, the more popular it is because there are more survivors".

[ii] Prognosis describes the likely disease course, irrespective of treatment. Prediction on the other hand describes the response to a specific treatment.

largely sought to consider tumours with greater molecular resolution, in effect making use of ever more abundant sources of data to more accurately classify a tumour with respect to treatment. However, it remains important to note that counterbalancing these attempts is a need for real-world clinical feasibility and accessibility.

1.3 Breast cancer classification

Attempts to realise classification and prediction in breast cancer have undergone a number of re-inventions, from class prediction based on anatomical location, through assessment of pathological presentation, immunohistochemical (IHC) molecular determination, to the current state of high-throughput multi-omic research. Each has improved patient stratification, though as-of-yet none has fully realised their potential in precisely assigning treatment. Key to the concept of patient stratification, is relating descriptive characteristics to clinical differences, be it response to a particular therapy or simply differences in survival. Though most early attempts relied on the anatomical location of the tumour, it has since become apparent that, although breast cancers may more closely resemble one another than they do gliomas, there remains considerable intra-class heterogeneity that location alone fails to capture, a paradigm that has defined contemporary cancer research.

Early breakthroughs in microscopy[i] allowed solid tumours to be defined beyond simply their location, by their clinical and histological presentation and imaging methods[ii] continue to be routine in breast cancer diagnosis. However, these methods are only suggestive of a lesion and histological confirmation is necessary, including determination of grade, size, vascularisation, lymphatic involvement and ability to invade surrounding tissues and organs. The World Health Organisation (WHO) continue to advise classification of invasive breast cancer based on these factors[14], and formal classification schemes, such as the Nottingham Prognostic Index (NPI), allow clinicians to systematically determine prognosis[15–18]. However, the variation in clinical course between patients with similar presentation remains high, accurate prediction of a response to therapy remains low and an ability to be mechanistically informative is lacking. Critically, histological grading is subjective and failure to reliably classify tumours highlights not only the difficulty posed by true tumour het-

[i] Johannes Müller and Rudolf Virchow first described the cellular origin of cancer by microscopy, moving away from esoteric descriptions of dysregulated blastema and the bodies four humors — blood, phlegm, yellow bile and black bile. Thomas Ramsden Ashworth would describe early tumour cells as “large and beautifully pellucid” and successfully isolate cells from a patient’s blood closely resembling those of the primary tumour, for the first time implicating the vascular system in metastatic spread. Alternative hypotheses had included “irritations” travelling along lymphatic vessel walls and the nervous system invoking the ‘idea’ of a tumour at a distal site. Ironically, an alterative hypothesis of spread by “morbid juices” proved to be closer to the truth.

[ii] Four major diagnostic screening methods are employed: clinical examination, mammography, ultrasound and magnetic resonance imaging.[13]

erogeneity but also in perceived heterogeneity, typified by modest rates of inter-observer agreement, and contention over the existence of some forms[19]. Tumour grading equally fails to highlight driving features that may be clinically actionable, describing effect rather than affect, and therefore remains a prognostic rather than predictive measure for hormone therapies.

Currently only three predictive markers are routinely used in the clinic: the oestrogen, progesterone and human epidermal growth factor receptors (ER, PR and HER2)[20–23]. Both oestrogen and the oestrogen receptor have been associated with invasive breast cancer for more than a century,[i] and these three molecules have largely defined breast cancer as at least three different diseases — loosely defined as ER+/PR+/HER2-, ER-/PR-/HER2+, and ER-/PR-/HER2-, though all permutations can present. Defining tumours by their driving molecular features has facilitated the use of targeted therapies, to disrupt some facet of the receptor signalling processes. In concert, these same molecules function as predictive biomarkers able to be quantified by IHC staining and allow treatment to be prescribed on a patient-by-patient basis. Unfortunately, however, whilst an absence of each of these markers preclude response to their corresponding therapies, presence only modestly correlates with response[25]. Moreover, of the approximate 60% of early stage breast cancer receiving adjuvant chemotherapy only, a fraction display benefit[26,27]. This limited response to therapy has fuelled an interest in whole-omic analysis, to transition from single molecule biomarkers towards — in theory — increasingly sensitive and descriptive multigene signatures.

1.4 High-throughput classification

The advent of gene expression microarray technology revolutionised our understanding of breast cancer as a disease of multiple transcriptionally distinct entities and influential 21st century studies continue to inform our understanding of its inherent heterogeneity[28,29].[ii] Importantly, whilst these newly defined subtypes correlated with IHC-based approaches, they additionally redefined seemingly homogeneous populations as composites of multiple sub-populations, each with their own profile of clinicopathological features, responses to therapy, and outcomes. A number of attempts to robustly define the true population level distribution and existence of potential molecular subtypes took place over the course of the early 21st century, though, like the IHC subtypes, they were again challenged by discordance

[i] George Beatson would demonstrate ovariectomy induced regression of the primary tumour in 1896, suggesting the role of steroid signalling as a major factor in breast cancer[24]

[ii] In fact two technologies for the total quantification of gene expression were published in the same 1995 edition of Science. Serial Analysis of Gene Expression (SAGE)[30] and cDNA comparative hybridisation (i.e. microarray chemistry)[31] were highly dissimilar in protocol and microarray-based approaches would prove easier to accomplish, particularly for high-throughput studies.

regarding their true existence and number. The most widely cited 'standard' for breast cancer subtypes remains Perou et al.'s[32] seminal class-discovery study. The group's hypothesis was simple, that the observable differences between tumours must relate to molecular differences, and that this molecular heterogeneity could be harnessed to stratify patients.[i] Most excitingly, these molecular differences, at the time understood to be driving the cancers, could potentially be determined in advance of any clinically observable features. Taking 65 unique invasive breast tumour samples, representing 42 individual patients, the group used newly available microarray technology to analyse transcripts corresponding to 8,102 genes, a significant proportion of the entire transcriptome. Hierarchical clustering grouped tumours and transcripts according to similarities in gene expression, illustrating the heterogeneous nature of breast cancer in a previously unseen clarity. Notably these differences were more often greater between patients of otherwise similar presentation than between samples from the same tumour following treatment. That treatment is a process fundamentally designed to disrupt a tumours normal state underlines the recurrent theme and challenge imposed by tumour heterogeneity. Pursuing a supervised approach, Perou and colleagues derived a subset of genes optimised for consistency between multiple samplings of the same tumour and variation between different tumours. This 'intrinsic' gene set revealed five robust tumour subclasses — Luminal A (LumA), Luminal B (LumB), HER2-positive (HER2), Basal-like and Normal-like. These would later be refined and shown to correlate with prognosis and response to therapy[34–36], as well as display distinct patterns of metastasis, with individual subtypes showing preference for particular organs. These were often shared, with bone metastases being common for all subtypes, but also distinct, with HER2-positive tumours more commonly metastasising to brain, liver, and lung tissues compared with ER-positive disease. Similarly, and again compared to Luminal A tumours, basal-like tumours displayed increased rates of brain, lung and distant nodal metastases, though significantly lower rates for bone and liver[37]. Despite this seeming clinical relevance, the 'intrinsic' subtypes would at first struggle to translate to clinical relevance[38], though HER2-positivity has since been incorporated clinically in both the PAMELA[39] and NOAH[40] trials, selecting for patients suitable for blockade with the monoclonal antibody trastuzumab. Interestingly, though the Normal-like subtype likely originated due to the inclusion of reduction mammoplasty in early subtyping studies, it continues to be observed in ~7.8% of lymph-node negative cases, exhibits poor outcome despite a shared IHC profile with Luminal A tumours[41] and has been conjectured to overlap with a subsequently defined subtype Claudlin-low[42].

A decade later, redefinition of the breast cancer subtypes was attempted using gene

[i] A proof of concept for this approach came from the Broad Institute's Todd Golub and Eric Lander who derived a gene signature whose expression could blindly classify acute myeloid and acute myeloid lymphoblastic leukaemias, including distinguishing B- and T-cell acute lymphoblastic leukaemia.[33]

copy-number alterations in ~2,000 patients, formalised as 10 'integrative clusters', ICs 1-10[43]. The integrative clusters exhibited distinct patterns of both survival and chemosensitivities and may therefore represent both prognostic and predictive stratification. Equally, they have been demonstrated to aid contextualisation of RNA sequencing-derived genomic drivers, certainly more so than the intrinsic subtypes[44]. However, as-of-yet, the IC10 classification scheme has failed to progress to clinical adoption, perhaps even hampered by its additional subtypes, where tumours are less confidently assigned to one subtype as a result.

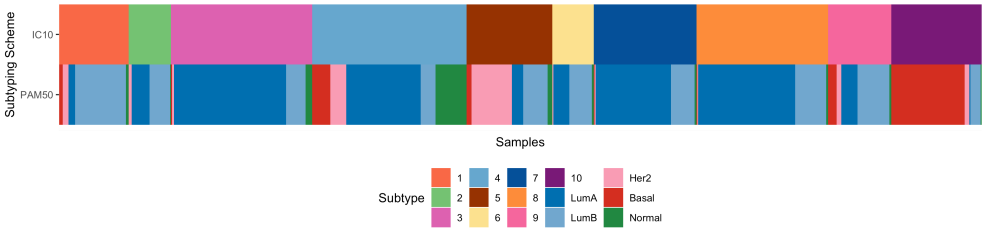

Figure 1.4.1: IC10 and intrinsic subtypes. IC10 subtyped patients are ordered left-to-right for ICs 1-10. The corresponding PAM50 subtype is shown below. LumA, LumB, HER2, Basal and Normal are coloured as dark blue, light blue, pink, red and green. Adapted from Ali *et al.* 2014.

Though molecular characterisation revealed and described the degree of breast cancer heterogeneity, to an extent molecular subtyping studies simply re-described the IHC subtypes. Even the ten integrative clusters display a high degree of synonymity with ER and HER2 statuses, with cluster 10 in particular overlapping with basal/triple-negative subclasses[44]. This was perhaps expected but also somewhat unfortunate as basal and triple-negative breast cancer (TNBC) subclasses remained defined by a lack of oestrogen and HER2 signalling, rather than any driving molecular feature. This has resulted in the lack of a TNBC-specific treatment, where patients continue to be offered chemotherapy alone, with almost all of the improvement in breast cancer mortality stemming from ER- and HER2-positive diseases. The potential of further unsupervised characterisation and subclassification was therefore questionable, though the need for targetable drivers and more reliable clinical stratification was clearly necessary. This would ultimately encourage a supervised approach, in an to attempt to define a direct molecular relationship with clinical outcome and treatment.

1.5 Prognostic characterisation

Whilst molecular class discovery had proven to be feasible, and at the time hinted at being mechanistically informative, they were at heart purely descriptive models.[i] To achieve clinically applicability, there was a need for clinical and molecular data to be considered in concert rather than in isolation. Demonstrating a degree of foresight, The Nederlands Kancer Instituut (NKI) had systematically collected frozen tumour samples with matched clinical histories since the early 1980s. Twenty years later, the availability of whole-transcriptome analysis offered the synthesis of these data to ask, not only which features define a tumour, but which defined its relationship with survival. Of these biobanked samples, a cohort of 78 node-negative women, allowed a supervised approach to reduce the available ~25,000 microarray probes to a signature of 70 genes correlated with the occurrence of a distant metastasis[49]. The team, led by Laura van't Veer, validated their findings in an additional collection of 295 breast tumour samples,[ii] where the 70-gene signature proved to outperform the then current clinically available predictors, including grade, nodal- and hormone receptor statuses and represented a conceptual validation of high-throughput molecular characterisation[50]. Further validation by the TRANSBIG consortium[51] again outperformed standard clinical assessment. The Amsterdam 70-gene signature would later be formalised as MammaPrint, the first FDA-approved multigene test, allowing prospective validation, in competition with AdjuvantOnline![iii], in the RASTER (n = 427)[52] and MINDACT (n = 6,693)[53] studies. Interestingly, of the 1,550 MINDACT patients discordantly classified as high clinical (AdjuvantOnline!) but low genomic (MammaPrint) risk and not receiving chemotherapy, 94.7% had at least 5-years distant metastasis-free survival (dmfs). These patients therefore represented a subset who would otherwise have needlessly been offered chemotherapy.

Encouraged by the potential for accurate determination of prognosis the search for additional prognostic gene signatures began in earnest. Despite the analytical and clinical validity demonstrated by MammaPrint, it's clinical utility was hampered by a reliance on microarray technology and the associated practical and economic costs. Of the subsequent generation of prognostic tests, Oncotype Dx, Prosigna, EndoPredict and the Genomic Grade Index (GGI) would stand out in satisfying the need for clinical and analytical validity as well as clinical feasibility.

Oncotype Dx measured a greatly reduced 21-gene signature by reverse transcription polymerase chain reaction (RT-PCR), returning a recurrence score (RS). Of the 21 genes,

[i] Subsequent work would later demonstrate that the unsupervised determination of the intrinsic subtypes did indeed correlated with response to treatment and prognosis[45–48]

[ii] In reality, 61 of the 295 patients included in the 70-gene signature validation were from the original training set, complicating the independence of the validation

[iii] AdjuvantOnline! allows online risk assessment based on clinical factors such as age, menopausal status, ER-positivity and nodal involvement amongst others.

16 were cancer-related non-housekeeping genes measuring ER- & HER2-signalling, proliferative and invasive potential. Patients scoring below 18 would be classified as low risk; above 30 as high risk. Oncotype DX would be retrospectively validated in 668 ER-positive and node-negative women, demonstrating a significant difference in 10-year distant recurrence rates for each RS category ($p < 0.001$), independent of age and tumour size[54]. Similar conclusions were demonstrated in 651 patients (including patients from the original training set) receiving adjuvant tamoxifen alone or in combination with chemotherapy, with a RS of >30 indicative of chemotherapy benefit[55]. Prospective validation as part of the TAILORx (Trial Assigning Individualised Options for treatment (Rx)) attempted to definitively determine chemotherapy benefit for recurrence scores of <10 but also whether that benefit was shared for intermediate scores of 11-25. The TAILORx study would demonstrate in 2015 that low RS participants share a very low risk of 5-year recurrence with endocrine therapy-alone[56]. More recently in 2018, 6,711 hormone-receptor positive, HER2-negative, and axillary node-negative patients scoring between 11 and 25 inclusive were compared by random assignment to either chemoendocrine therapy or endocrine therapy alone. Endocrine therapy alone was found to be noninferior to chemoendocrine therapy for all end points investigated, suggesting that chemotherapy in addition to endocrine therapy for intermediate recurrence scores and below could also be avoided, representing up to 85% of women with early breast cancers[57]. Though originally intended for use in node-negative disease, recurrence scores have also been demonstrated applicable in women with up to three tumour-involved axillary lymph nodes[58].

Refinement of the intrinsic subtypes as a 50-gene signature (PAM50)[35] allowed for FDA approval and commercialisation in post-menopausal and HR-positive tumours as the Prosigna Breast Cancer Gene Signature Assay. Prosigna would use NanoString technology, a microarray-like alternative, to accurately and efficiently measure signature transcript abundance. Intrinsic subtype classification and calculation of a risk of recurrence (ROR) score that, like to Oncotype DX, assigns low, intermediate and high distant recurrence risk categories on a 0-100 scale. Validation in 786 ER-positive women of mixed nodal status receiving adjuvant tamoxifen compared ROR score, IHC and clinical factors, demonstrating Prosigna's improved ability to predict disease-specific and recurrence-free survivals[59]. Further validation comparing 1,478 post-menopausal ER-positive women receiving endocrine therapy-alone again demonstrated accurate recurrence-free survival and prognostic improvement over conventional methods[60].

An eleven gene PCR based signature of endocrine treated ER+/HER2- 10-year recurrence risk — EndoPredict — would show similar promise in training and validation studies[61,62]. It would be integration with clinical factors, as EPclin, however, that would demonstrate marked improvement over standard methodology, re-predicting a subset of high-risk pa-

tients as having favourable prognosis[63]. Further validation in 555 chemotherapy treated FFPE (formalin-fixed and parafin embedded) samples would establish EndoPredict and EPclin's utility in independently predicting metastasis-free survival[64].

Finally, an alternative approach was taken by Sotiriou et al.[65], to use microarray-derived expression of 97-genes to directly infer grade, itself associated with prognosis, as the Genomic Grade Index (GGI). Comparing differentially expressed genes in 64 ER-positive breast tumour samples, the GGI allowed partitioning of patients in to 1 of 2 categories — low or high — which roughly corresponded to grades 1 and 3 respectively. Importantly, the GGI does away with the intermediate IHC-derived grade 2 category, where patient tumours are assigned as either of the GGI's own 2 categories. The GGI was validated in 597 breast cancer samples, attempting to find an association between GGI and recurrence-free survival, and indeed high GGI was associated with a higher risk of recurrence.

It was clear that clinically feasible and prognostic gene signatures were of relevance in breast cancer treatment, albeit predominantly for ER-positive disease. To compare *between* signatures would prove challenging however, and determination of a one-size-fits-all model was proving unlikely. For any given signature, it was likely that its use would roughly extend only to the context of its development, that is: the response measured (clinical, histological, biological); the cohort studied (subtype specificity, pre- vs. post-treatment); and the treatment used (chemo- and endocrine therapies). Despite these difficulties, Fan et al.[66] would compared 5 signatures, including MammaPrint, Oncotype Dx and Prosigna, using the 295 sample microarray-measured MammaPrint validation dataset. Whilst not identical, each individual signature largely demonstrated concordance in assignment. Perhaps surprisingly MammaPrint and Oncotype Dx would show improved concordance when considering all samples (81%), opposed to ER-positive only tumours (77%), despite Oncotype Dx's development using strictly ER-positive datasets.

1.6 Lessons from high-throughput gene expression studies

A wealth of data was now available to researchers world wide. As the community called for open-access to datasets, particularly those derived from patient-donated material, journals began necessitating that researchers deposit data in publicly-accessible databases, allowing for further/re-analysis and study validation. The fulfilment of microarray-derived data's promise appeared inevitable. However, as more studies reported prognostic signatures, it became clear that their ability to outperform current methods was underwhelming, that the overlap between signatures was small and, moreover, the total proportion of the genome being reported was larger than anticipated. This represented a conundrum: how could the diversity of prognostic signatures all be correlated with disease outcome and, by

extension, with their underlying mechanisms?

Perhaps the most obvious (and troublesome) explanation was that microarrays were simply non-reproducible and subject to non-specific hybridisation or, put another way, were completely unreliable. To challenge this, a consortium effort by the Microarray Quality Control (MAQC) consortium assessed microarray reliability, performing a series of systematic reproductions across multiple different platforms and laboratories, demonstrating that not only was there high reproducibility but also confidence for clinical application[i][67,68]. Another possible explanation was hinted at by the unintended synonymous first-generation signatures describing the ER-positive/ER-negative axis. In effect, because these subtypes differed so consistently in their clinical outcomes, by comparing favourable and poor outcome patients researchers were in essence comparing ER-positive to ER-negative disease by proxy, and deriving genesets that re-described this relationship[69]. The diversity of equally prognostic signatures could therefore be explained as a large pool of highly correlated members of the same molecular pathways, as is the case for ER-target genes[70]. These correlated genesets would be able to be substituted for one another with little effect on the signature's overall prognostic capacity[71]. Understanding of the interrelatedness of clinical, molecular and prognostic features would make this result all but inevitable in retrospect, and that many of the derived genesets were often only as effective as IHC ER measurement was therefore unsurprising.[ii]

It was clear that assuming any given gene was unlikely to be associated with prognosis was misguided. Venet et al., would go one step further demonstrating that most signatures were no more prognostic than *random* subsets of genes, and that any single gene was reasonably (5-17%) likely to be associated with prognosis[72].[iii] The explanation proposed was that a large proportion (>50%) of the genome correlated with proliferation and that, similar to ER target genes, any number of genes could be substituted for one another with little effect on overall prognostic association. This was characterised as meta-PCNA of which correction for could alleviate a seemingly spurious association with prognosis. However, it would later be demonstrated that random signature association did not necessarily extend to all cancer types, despite their shared proliferative nature[73]. Again, as grade was routinely measure by histological methods, it is no surprise that these signatures of proliferative capacity failed to decidedly outperform existing methods.

[i] However, though reproducibility had been demonstrated, the use of high-quality RNA across the MAQC experiments was nonetheless not wholly representative of early gene expression signatures derived from FFPE material.

[ii] Salomon saith, There is no new thing upon the earth. So that as Plato had an imagination, that all knowledge was but rememberance; so Salomon giveth his sentence, that all novelty is but oblivion (Francis Bacon: Essays LVIII).

[iii] That this study was performed using the dataset originally used to derive 70-gene MammaPrint signature was particularly worrying.

It became clear that to avoid simply restating the ER+/- axis and/or measuring proliferation, more homogeneous patient sub-populations must be analysed. Iwamoto and Pusztai[71] would comment

> "Investigators who developed the first generation of supervised prognostic and treatment response predictors started with the then prevailing notion that breast cancer is a single disease, and all subtypes of breast cancer were included in the analysis. This resulted in major limitations in the diagnostic products that emerged from this research."

However, stratifying cohorts would itself prove difficult, particularly for low-prevalence subtypes. By grouping patients prior to an analysis, the sample size would inevitably decrease and, by extension, the statistical validity. By pooling data from several sources, however, researchers hoped to circumvent this pitfall, though, again, a naive assumption of consistency would prove a stumbling block. Even data processed in the same laboratory as part of sequential batches was liable to show systematic technical (i.e. non-biological) variation and would therefore require normalisation before comparison[74]. Correction methods would prove difficult to engineer, needing to remove technical batch differences alone, leaving true biological variation unaffected. Ultimately, minimal correction of only suitably similar datasets would be encouraged, though the definition of suitable similarity remains ill-defined, as does the balance between improving sample size whilst respecting dataset homogeneity. Complicating matters still further is the observation that even when stratified by ER, some breast tumours remain inherently difficult to classify, being consistently incorrectly classified by nearly all available signatures[69]. Moreover, analysis of ER-positive cohorts simply tend to learn HER2 status and *vice versa*.

The challenges inherent in study design and cohort selection were also apparent in the extent to which prognosis could be calculated. As almost all studies used in the design of prognostic classifiers were developed using cohorts that had ~10-years follow-up, these classifiers understandably have a 10-year cut-off for what they can predict. Whilst partially reflecting the realities of cancer survivorship, it is equally true that recurrences do occur well beyond this threshold and remain uncharacterised. Equally by considering patients as a cohort, application remained relevant to this group and suffered limitations in predicting the outcome of an individual. In fact, despite all the research that has gone into prognostic and predictive classifiers, clinically the only accepted predictive measures for breast cancer remain ER and HER2 by IHC.

It could be argued that microarrays have failed to live up to their early promise in clinically impacting cancer and, largely, the field has moved on to sequencing based methods of-

fering the opportunity to explore novel as well as known transcripts and their isoforms, as well as single-cell approaches[75]. It should, however, be stressed that the microarray era has been instrumental in furthering our understanding of the inherent heterogeneity of breast cancer; the comparable relationship between metastatic and primary tumours[76,77]; the fundamental role of proliferation, mitosis and chromosomal integrity in aggressiveness[78,79]; and the interaction between tumour and stromal compartments[80,81]. Venet et al. themselves were cautious to note that the shared prognostic potential of derived and random signatures did not invalidate that potential, further commenting that prognostic confounding extended to pre-genomic era studies[72].

Though progress towards predictive biomarkers has been piecemeal this understanding has and continues to inform future studies. Furthermore, as the potential to combine and interrogate pre-existing datasets improves, it remains to avoid throwing the baby out with the bathwater. The wealth of microarray data available in the public domain remains highly informative, if suitable questions are asked of it, and it remains a cost-effective assay to maximise sample size. Ultimately, careful independent validation across multiple data sources is critical to remain confident that the variation and heterogeneity measured is a factor of the disease and not the dataset.

1.7 Tumour heterogeneity and evolution

The impact of heterogeneity is extensive, and in the context of cancer our consideration of heterogeneity is twofold as (i) spatial heterogeneity at a distinct moment in time and (ii) temporal heterogeneity, the evolution of a tumour over time. Importantly, both spatial and temporal heterogeneity can themselves be subclassified as *inter*- and *intra*- tumour heterogeneity. Inter-tumoural heterogeneity is simply the variation between different individuals tumours (or between multiple tumours from the same individual), and intra-tumour heterogeneity instead defines the variation inherent within a single tumour mass, either at a single point in time or over time as multiple observations compared against themselves. We can consider response (or lack of response) to treatment as a facet of temporal heterogeneity and, in many senses, temporal heterogeneity is the goal of treatment, to chemically coerce a tumour towards a less aggressive and invasive phenotype over time. However, heterogeneity more often than not complicates matters, both experimentally and clinically, where any observation inevitably represents a cross-section in both time and space and is at risk of not fully representing a dynamic tumour[82–84]. Given that treatment is prescribed based on these snapshots reveals a diagnostic failing — that correct classification of a biopsy does not guarantee correct classification of the tumour(s).

Key to the concept of temporal tumour heterogeneity is the well characterised sequen-

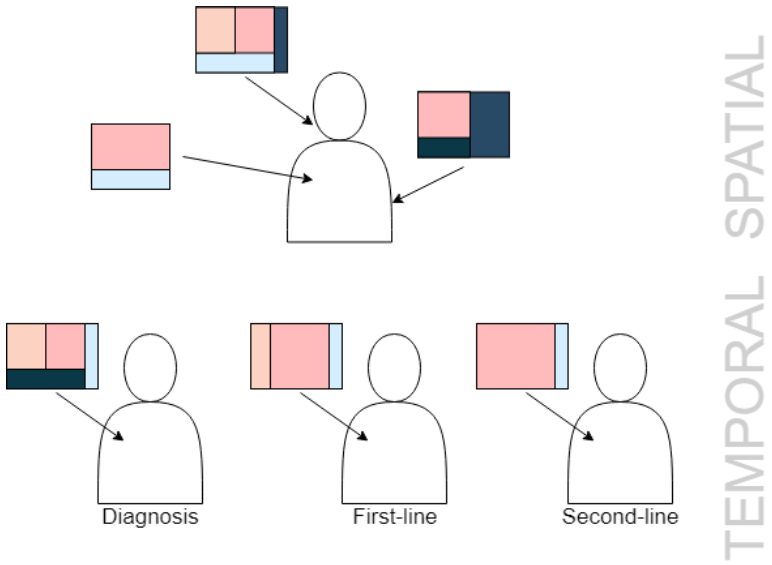

Figure 1.7.1: Breast cancer heterogeneity. Heterogeneity can manifest in two ways, with variation by either position or by time. For temporal heterogeneity, tumour subclones are seen to vary in response to treatment.

tial acquisition of genomic lesions and alterations that additively destabilise cellular function and drive the cancer phenotype[85,86]. Classically this has been presented as a linear development[i] but more recently it has become clear that, following a stochastic initiating event, multiple subclonal populations are able to compete and evolve in tandem, expanding or diminishing dependant on their environmental fitness[88].[ii] This process of evolution is consistent throughout the lifetime of the cancer, extending beyond its initial establishment and allowing a dynamic response to any destabilising changes in its environment, most pertinently in resistance to therapy[90,91]. As fitter clones emerge, their potential to vary likely increases, allowing the tumour to rapidly acquire fitness under changing circumstances, highlighting the difficulty in addressing highly aggressive and resistant metastases and the need to adequately profile and treat a tumour as thoroughly and early as possible. Even the concept of "early" is challenging in this regard, and metastatic potential is likely to be evident even at the clonal stage of the primary tumour[76].

The implications of heterogeneity therefore include variation between ostensibly similar separate tumours (inter-tumour) as well as within a single tumour housing subclonal regions (intra-tumour), any of which may respond differently in the clinical setting. Whilst, inter-tumour heterogeneity is likely a combination of germline and accrued somatic mutations in addition to environmental factors, intra-tumour heterogeneity is understandably fuelled by somatic alterations, as any present tumour subclones will share identical

[i] And is still representative of haematological malignancies.[87]

[ii] One hypothesis reads that, in larger mammals, clonal diversity becomes a problem for the tumour itself. As heterogeneity increases, the likelihood of an ultra-aggressive clone developing also increases. This *hypertumour* effectively becomes 'cancerous' to the original tumour, destroying it.[89]

germline and macro-environmental starting points. Though whole-genome single-cell sequencing has demonstrated that any given tumour cell is likely to be genetically unique[92], from a clinical perspective the role of tumour heterogeneity needs only to be as complex as that which directly impacts upon patient survival and well-being.[i] Attempts to disentangle non-specific from key driver alterations has proven difficult, particularly in breast cancer, with no single genetic mutation proving integral to tumour function; thereby limiting the ability to selectively target an 'Achilles heel'. Instead many cancers are diverse in their driver mutations, with no one mutation being consistently found in more than even 20% of 1,013 prostate tumours[94].

Clonal evolution in response to therapy, including targeted agents, is dynamic, unpredictable and largely synonymous with resistance[95–97]. Selection is able to occur during the course of treatment, with the EGFR tyrosine kinase inhibitor erlotinib inducing expansion of resistant $EGFR^{T790M}$ mutant clones in a lung cancer patient, but also in its absence, with these same resistant clones diminishing and becoming undetectable following a 10-month treatment-free period[95]. This highlights that the most resistant clone may not necessarily represent the broadest fitness, potentially only able to flourish under the extreme selective pressure of a targeted therapy. It has therefore proven beneficial to perform longitudinal tumour sampling, with a focus on observing a response *to* treatment, rather than predicting from a static pre-treatment or surgical biopsy alone. Routinely, where multiple tumour biopsies have been collected, diagnostic core-needle biopsies (CB) and surgical excision biopsies (EB) have been used to compare pre- and post-treatment effects. To an extent this has been limited to a small number of pre-defined molecular markers, including the routinely measured ER, PR, HER2 and Ki67, though whole-omic investigations have also been performed, investigating the effects of anthracycline-based chemotherapy; celecoxib; RAD001; anastrozole; letrozole; exemestane; and fulvesterant[98–107].

Ultimately, heterogeneity increases treatment complexity and the tumour represents a moving target. Even for ER-positive disease, initial or eventual resistance to endocrine therapy is commonplace and it is in characterising and predicting these patients that chapter 4 of thesis will focus.

1.8 Endocrine therapy

For a patient presenting with ER-positive disease, several systemic treatment options (summarised Table 1.1) are available, including chemo- and endocrine therapies. Endocrine therapy specifically targets susceptibilities in oestrogen dependent tumours at the expense of much of the toxicity associated with chemotherapy[108]. Moreover, this decreased toxi-

[i] And in fact heterogeneity in of itself is a prognostic marker.[93]

city has little bearing on their efficacy, with several studies reporting greater benefit in the adjuvant setting for post-menopausal women[109]. By disrupting oestrogen biosynthesis or signalling, endocrine therapies achieve cell cycle arrest in G1/S, and are therefore non-cytotoxic[110]. Selective oestrogen receptor modulators (SERMs, including the widely pre-menopausally prescribed Tamoxifen) antagonistically[i] compete for intra-nuclear ER, inducing a conformational change and destabilising its ability to bind oestrogen response elements (EREs). However, this inhibitory effect on oestrogen signalling is incomplete, with a partial response still evident[113]. Selective oestrogen receptor downregulators (SERDs) also bind competitively to dimerised ER but instead facilitate downregulation by degradation, avoiding any residual partial oestrogen agonism[113,114]. In the post-menopausal setting, aromatase inhibitors (AIs) act to destabilise the aromatase enzyme (CYP19), downregulating oestrogen biosynthesis occurring in subcutaneous fat stores prevalent in peripheral and breast tissue[115]. Aromatase catalyses androgens as part of the final step in producing oestrogens, where it can be irreversibly and inactivatingly (Type I AIs), or reversibly and competitively (Type II AIs), inhibited[116,117]. Results from the ATAC (Arimidex, Tamoxifen alone or combined) trial[118] have shown anastrozole to outperform tamoxifen in the post-menopausal setting and improve disease-free but not overall survivals in the adjuvant setting.

Endocrine treatment selection is based on toxicity, tolerability, the potential for partial ER agonism, and menopausal status[119]. In the post-menopausal setting, results from the FIRST (fulvestrant first-line study comparing endocrine treatments) trial have suggested fulvestrant to be equally as effective as anastrozole as a first-line therapy but with an increased time-to-progression and overall survival, as well as increased progression-free survival as a second-line therapy[120,121].

SERMs	SERDs	SERM/SERD Hybrids	AIs
Tamoxifen (1977)	Fulvestrant (2002)	Elacestrant (Phase II)	Anastrozole (Type II, 1995)
Toremifene (1997)			Letrozole (Type II, 2005)
Raloxifene (2007)			Exemestane (Type I, 2011)

Table 1.1: Endocrine therapies and FDA approval dates.

The degree of choice and specificity of application highlights the difficulties in prescribing treatment, even within the seemingly homogeneous collection of patients exhibiting ER-positivity. Further complicating matters is that correct prediction of a response to endocrine therapy is only half the battle, and histologically confirmed ER-positive tumours are observed to display a lack of response to endocrine therapies in 50-70% of cases[104,122]. It should be noted that this number ignores the fraction of initially responding patients who

[i] Tamoxifen conversely agonises ER in the uterus and liver, as well as demonstrating species-specific effects[111]. Agonistic action is thought to explain small increases in the rate of endometrial cancers.[112]

eventually become desensitised to treatment, an inherently difficult dynamic to study experimentally. Taken together, these highlight the clear benefit that could be derived from improved models of endocrine resistance, not least in producing predictive biomarkers of response to endocrine therapy.

1.9 Endocrine therapy resistance

Resistance to endocrine therapy is pervasive in ER-positive disease though, broadly, resistance mechanisms are common for both targeted and chemotherapies. Superficially, these can be defined by drug target alterations, activation of pro-survival compensatory mechanisms, and avoidance of programmed cell death.[i] Underpinning these escapes are both specific alterations, including mutations in — or overexpression of — the drug target and/or regulatory elements and post-transcriptional modifications; as well as more general changes at the level of the cell, including DNA damage repair pathways, alternative signalling activation and non-specific drug efflux. Within the context of ER-positive disease, multiple mechanisms of resistance to endocrine therapy have been proposed, though a definitive explanation remains inconclusive. Perhaps surprisingly, loss of ER expression is not common amongst these, with only ~10% of initially ER-positive tumours later presenting as ER-negative[123]. ER mutants often remain sensitive to second and third line antiestrogenics, targeting varying facets of oestrogen signalling, though at iteratively reduced rates[124]. Further complicating the role of mutant ER in endocrine resistance is evidence of heterodimerisation with wildtype ER. Nonetheless ER mutants have been functionally characterised *in vitro*[125–127] and shown to be enriched in the metastatic lesion compared to the primary tumour[128–132].

More commonly, dysregulation downstream of ER is observed, including interactions between ER coactivators and corepressors, and alternative signalling molecules. Ligand independent ER phosphorylation and activation via growth factor signalling pathways, such as insulin/IGF (insulin-like growth factor) and epidermal growth factor receptor (EGFR) superfamilies[133,134] converge on mitogen-activated protein kinase (MAPK) family kinases able to activate ER by phosphorylation. Ligand independence represents a possible escape from aromatase inhibition, as well as in activating parallel programs of growth and survival[135]. Though cross-talk between ER signalling and growth-factor receptors has been described, there is little evidence that combination therapies can counteract the resistant phenotype[101,136–138], potentially reflecting the relative deficiency of cell line models in investigating the vastly complex nature of cancer and resistance. A notable potential exception

[i] Here we will only discuss pharmacodynamic mechanisms of resistance, though pharmacokinetic mechanisms (variation in drug absorption, distribution, metabolism and elimination) are also factors liable to effect drug responsiveness.

to this is the PI3K/AKT/mTOR signalling pathway, for which a number of targeted therapies are currently in clinical trials. PI3K/AKT/mTOR is frequently mutated in breast cancer[139], and has been demonstrated to impart AI resistance in cell line models[140]. FDA approved everolimus, a TORC1 inhibitor has shown some promise in combination with chemotherapy (exemestane) or tamoxifen for patients with advanced ER-positive disease[141]. However, inhibition of PI3K/AKT/mTOR pathway members has alternatively been demonstrated to remove the pathway's own inhibitory checks as well as activate pro-survival factors[142–145].

Resistance mechanisms independent of oestrogen signalling include alterations to tamoxifen metabolism by *CYP2D6*, which converts tamoxifen to antiestrogenic endoxifen and 4-hydroxytamoxifen[146]. Allelic variation in *CYP2D6* impacts upon its ability to adequately metabolise tamoxifen and, therefore, poor metabolism presents both a mechanism of resistance as well as a potential predictive genotype[115].

Resistance can be realised in two ways: *de novo*/intrinsic (existing prior to treatment) or acquired resistance. Intrinsic resistance implies that a pre-existing resistance mechanism is apparent in the tumour at diagnosis and that tumour is immediately insensitive to treatment. Acquired resistance occurs in response *to* treatment, where a previously responsive tumour eventually fails to respond. For ER-positive disease desensitisation to first-line therapy and acquisition of resistance is evident in up to 30% of cases[148,149]. Second- and third-line therapies remain effective, though again at gradually reduced rates decreasing from a response in 70% of patients to only 30%[121,150]. Mechanistically, intrinsic and acquired resistances differ, and acquired resistance is associated with a wider mechanistic diversity[148], and this diversity is both patient as well as treatment specific[151]. It is worth noting that, though they develop over the course of treatment, the mechanisms of acquired resistance may be present within the tumour at diagnosis. For instance, a subclone of a heterogeneous tumour may be intrinsically resistant to therapy but it is only after the selective pressure of treatment that it is able to progress.[i] Whether intrinsic and acquired resistance are independent of one another remains debated and it is possible that the immediacy of resistance is simply a factor of the subclonal population of the tumour at diagnosis. The actual mechanisms of resistance may be shared between *de novo* and acquired instances and simply differ in the time of their presentation[152].

Together, this portrays a rather jumbled picture of multiple competing mechanisms and a paucity of patient-derived models. Acquired resistance in particular presents a research challenge due to its latent presentation, with ER-positive tumours displaying a steady rate of recurrence even decades post-surgery[153]. Key to this difficulty is in procuring research material — for any latently presenting phenotype, inevitably the tumour material has been

[i] His arrows they are only claws, his wings a pair of lies, his horns are hidden by the wreath, he is, we must surmise, like all the Gods of ancient Greece, a devil in disguise.

removed at the point of surgery and treatment. Acquired resistance has instead been characterised by cell line and animal models, typified by epithelial to mesenchymal transition (EMT) and an increased invasive and migratory phenotype within this context[154–157]. Equally, enrichment for stem-like progenitor cells in tamoxifen-resistant breast cancers has been linked with a mechanism for acquired resistance[158]. Specifically, Sansone et al. demonstrated subclonal upregulation of IL6 in $CD133^{hi}$ cells, in response to chronic ER downregulation in a patient-derived xenograft (PDX) model[159]. Though compensatory growth factor signalling has been proposed as a mechanism of sustained growth in the face of endocrine therapy[160], attempts to target these pathways have so far been unsuccessful and short-lived[161,162]. Furthermore, of the targeted therapies demonstrating most promise (CDK4/6 and mTOR inhibitors), there are associated toxicities that complicate administration and require accurate predictive biomarkers to be developed[163].

Key to acquired resistance is the long-term survival of residual disease, able to perpetuate in a true quiescent or balanced-turnover dormant state. This may manifest as only a single disseminated cell but importantly, time until clinical presentation is independent of cell doubling rates[164]. Instead cells can remain dormant for extended periods of time, upwards of 5-years, before they reawaken and begin to propagate. Characterising and predicting when this reawakening may occurs has been understandably challenging and again cell and animal models have made up the bulk of the literature on breast cancer dormancy. Current theories implicate escape from immune surveillance and/or microenvironmental effects or increased angiogenesis[165,166], though again these results are based on cell and animal models that may not translate and the need for patient-derived models of acquired resistance and dormancy are needed.

1.10 The adjuvant and neoadjuvant settings

Currently, five major treatment practices are used in managing breast cancers: surgery, radiotherapy, chemotherapy, hormone therapy and immunotherapy. Of these five, for the majority of patients surgery is associated with the largest single benefit[167], with the remaining four being employed as secondary, adjuvant measures. By considering clinical presentation of the disease and patient, breast cancer management has traditionally progressed as surgery followed by one or more adjuvant treatment — most commonly radio- and chemotherapies — to eradicate residual primary disease or micrometastases. Systemic therapies, such as chemo-, hormone and immunotherapies are now regularly given in an attempt to remove residual disease burden that remains post-surgery, either at the surgical site or as a metastasis. The success of this strategy is unequivocal and, for radiotherapy, adjuvant use following breast-conserving surgery has been shown to be as effective as total

mastectomy combined with axillary lymph node dissection[167]. However, due to the nature of the adjuvant setting, determining drug effectiveness as part of a clinical trial is likely to exceed 10 years. This is due to the fact that the most immediate read-out of response (i.e. whether tumour cells are being killed) has been removed by the act of surgery. This effectively leaves late stage events, such as relapse and/or metastasis, as the only viable surrogate measures, each of which can take decades to present clinically.[i] Beyond the obvious time limiting nature of this approach, patient monitoring and drop-out also present serious hurdles, affecting the statistical validity of any given study. Combined with the well-described attrition rates and long development time-scales for novel targeted therapies, this delay represents a significant obstacle to drug availability.

In response to these difficulties, and in an attempt to expedite clinical research, it was proposed that the pre-operative setting could instead be used to directly correlate treatment with a measurable affect on the *in situ* tumour. This neoadjuvant setting opened up a number of possibilities in improving a patient's clinical management. Ultrasound sonography (USS), mammography and/or physical calliper measurement were able to measure dynamic changes in tumour volume over time in response to treatment, revealing the success or failure of a prescribed treatment.[ii] If a treatment was observed to be ineffective, a second regimen could then be employed. As this treatment would be continued into the post-surgical adjuvant setting, treatment could be effectively tested on a patient-by-patient basis and, furthermore, any reduction in tumour volume would allow for less invasive surgery and increased the likelihood of breast conserving surgeries, potentially removing the need for surgery entirely[169].

It should be noted that neoadjuvant interrogation is dependent on accurate characterisation at the phenotypic level, most commonly in the measurement of tumour volume, used as a surrogate for response to therapy, and performed by the methods listed above. Like any method performed physically by a clinician, and relying on subjective assessment of the data in real time, technical variability is guaranteed. Moreover the technologies themselves are known to perform only within a tolerance of error, though USS measurement has been shown to outperform calliper measurement as well as demonstrate high re-test agreement as well as inter-observer agreement[170–172]. The practice has been used to successfully guide patient classification in a number of studies, including in a successful blind validation of endocrine therapy response prediction[102]. Conversely it has been demonstrated that alternative — though more involved — techniques such as CT (computed tomography) scans

[i] It would take 15 years (1974-1989) following the early adoption of the adjuvant setting to reliably prove that tamoxifen reduced breast cancer incidence.[168]

[ii] In fact a side-effect of improvement in breast cancer management is that long-term events, such as recurrence or metastasis, become less prevalent and statistical comparison become underpowered. The neoadjuvant window allows surrogate outcomes such as complete pathological response to be utilised in their place.

offer greater sensitivity and positive predictive value, at a cost of efficiency[173]. Furthremore, the determination of tumour size is potentially overestimated by USS, compared with CT[174].

Phenotypic classification often problematically involves hard divisions to be drawn, for instance defining responsive vs. non-responsive patients, where in truth a spectrum of response may exist. Here, again, Turnbull et al.[102] would demonstrate that a quick response to therapy, defined via rapid tumour volume reduction, and a slower response were transcriptomically comparable, encouraging classification despite the likelihood of shades of grey. It also remains that volume reduction of a tumour may liberate resistant clones that were otherwise kept in check, as discussed in Section 1.7, implying that tumour shrinkage may not always necessarily correlate with better prognosis in the long-term. It is critical to remember, however, that these methods are understood to be imperfect, and it is this imperfection that necessitates their use in clinical and molecular studies, in an attempt to derive more accurate and sensitive methods.

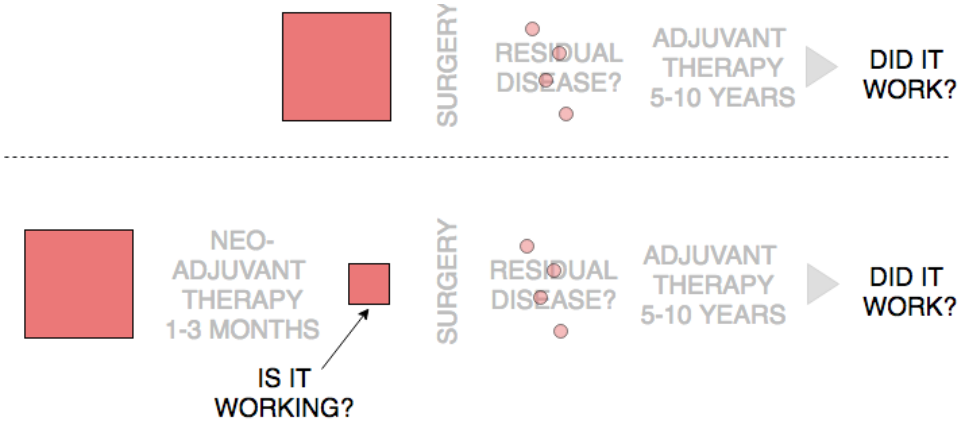

Figure 1.10.1: The adjuvant (top) and neoadjuvant (bottom) settings. Study times are often more than 5 years in length and adjuvant therapies can only be validated after this time has elapsed. By tracking the tumour prior to surgical excision, a dynamic response to therapy can be evaluated within a greatly reduced time frame and multiple biopsies can be taken from the *in situ* tumour.

From a research perspective, the neoadjuvant window offers the ability to collect tumour samples over time from the same patient. Multiple biopsies from the same patient allow an additional dimension to be considered in the dynamic molecular response to treatment — i.e. it's no longer just that the expression of gene *X* was high at diagnosis, rather that it has shown a significant reduction following treatment. By focussing on *changes* in the molecular profile of a tumour, researchers are able to add patient-specific context to their response or progression and effectively model temporal heterogeneity[102].

It should be noted that neoajduvant interrogation was dependent on accurate characterisation at the phenotypic level, most notably in the measurement of tumour volume,

used as a surrogate for response to therapy, performed by the methods listed above. Like any method performed physically by a physician or clinician, and relying on subjective assessment of the data in real time, variability is guaranteed to be apparent. Moreover the technologies themselves are known to perform within a tolerance of error [ref]. However, the practice has been used to successfully guide patient classification in a number of studies, including in a successful blind validation of endocrine response prediction[102]. Phenotypic classification, by USS for instance, somewhat problematically involves hard classifications to be drawn, often defining responsive vs. non-responsive patients, where in truth a spectrum of response may exist. Here, again, Turnbull et al.[102] would demonstrate that quick response, defined via rapid tumour volume reduction, and slower response were transcriptomically comparable, encouraging classification despite the likelihood of shades of grey. It also remains that volume reduction of a tumour may liberate resistant clones that were otherwise kept in check, as discussed in the previous chapters, implying that tumour shrinkage may not always necessarily correlate with better prognosis. It is critical to remember, however, that these methods are understood to be imperfect, and it is this imperfection that necessitates their use in clinical and molecular studies, in an attempt to derive more accurate and sensitive methods.

This has most notably been demonstrated as part of the IMPACT (immediate preoperative anastrozole, tamoxifen, or combined with tamoxifen), P024 (letrozole vs. tamoxifen) and POETIC (peri-operative endocrine therapy for individualising care) trials. IMPACT and P024 both independently demonstrated that Ki67 reduction between study arms was analogous to rates of recurrence observed in the ATAC trial[175,176]. However, key to utilising molecular features as surrogates for pathological outcomes, and symptomatic of a gradual growing scepticism for seemingly inconsistent high-throughput technologies, was a need to robustly and reliably determine biomarker consistency over time. A host of studies undertook this challenge, measuring hormone and growth receptors and Ki67 expression consistency by IHC, both in lieu of treatment[177] and under conditions of treatment[177–179]. One meta-analysis summarising these consistencies reported reliably accurate and concordant determination of ER, PR and HER2 statuses for both core and excision biopsies[180], though another described a meaningful difference in the determination of PR in particular[181].

That discordancy between core and excision biopsies existed was expected and the characteristics of change were intuitive. That large, aggressive and node-positive tumours are liable to show the greatest discordancy in Ki67, a marker of proliferation, following chemotherapy was self-evident but raised the question: does treatment-induced heterogeneity causing a loss of ER/Ki67 imply that treatment should be altered? Kim et al.[182] demonstrated that extreme Ki67 discordance between CB and EB measurements are correlated with poor prognostic factors, including size, PR-negativity, grade III, and early disease

onset. It has additionally been demonstrated that 12-21% of luminal A tumours defined by IHC at diagnosis would otherwise have been classified as luminal B if alternatively defined at surgery[183,184].

The spectre of a tumour changing subtype over time is a troubling one. Though treatment is routinely altered and fine-tuned throughout a patient's therapy, for instance in the case of resistance to anti-estrogenics, as discussed in "Endocrine therapy resistance" (Section 1.9), little has been written on how to deal with tumours losing a marker that was previously apparent. Indeed, simply monitoring changes in the primary tumour's subtype have only recently been possible within the framework of the neoadjuvant window. If we consider the situation where tumour subtype alters from ER-positive to ER-negative between treatment initiation and surgery, two contrasting possibilities are apparent — that treatment has effectively eradicated its particular target and should be continued, or that a once successful treatment no longer has a population that is sensitive to its effects. From a patient's perspective, the latter is particularly troubling, where the potential for useless overtreatment becomes increasingly possible. Effectively, as subtype changes the amount of data that can be used to assign treatment increases, where one, be it the original or the revised assignment, may better reflect the post-surgical prognosis. This is further complicated in that it is not necessarily routine practice to biopsy even recurrent disease when it presents. To date, most studies have only gone as far as characterising the potential for subtype discordance, without making any recomendation on how to manage this occurence. Where centres have altered their treatment in response to changing subtype, no survival analysis has yet been calculated[185]. However, there is evidence that discordancy is associated with a decrease in overall and post-recurrent survivals[186], and in the case of ER, differential loss following neoadjuvant chemotherapy has been linked to worse disease-free and overall survivals in a multivariate analysis[187], whilst the opposite appeared true for Ki67[188]. Despite these varying effects, results from the IMPACT trial confirmed Ki67 reduction as predictive of long-term survival in hormone receptor positive tumours, additionally suggesting that longer-term changes in expression may be observable at as few as 2-weeks on treatment, offering the potential for clinically feasible testing within the context of the neoadjuvant window[189,190].

Whilst treatment was evident to molecularly alter tumours and that these changes were associated with survival, several groups additionally suggested that the act of the tumour biopsy may directly impact upon both the data generated but also, more worryingly, the disease trajectory[191,192]. Surgical excision — and induction of a wound-healing state — of HER2-positive cancers has been implicated in promoting malignancy via systemic growth factor release and immune suppression[193,194]. Tagliabue et al. would further demonstrate that residual disease removal, effectively a second surgery, for HER2-positive tumours re-

sulted in a significant increase in proliferation[195]. Similarly, Chen et al. demonstrated a correlation between molecular subtype, surgical biopsy interval and Ki67, implying that the structure of the study was able to effect the data observed, specifically implicating the CB in upregulating Ki67[196].

Fundamentally confounding these studies was that the act of sample biopsy was concurrent with treatment — the effects of one were difficult to disentangle from the other, and it was therefore necessary to molecularly define changes occurring in lieu of treatment, particularly in instances where, as suggested by the IMPACT trial, molecular markers such as Ki67 were seemingly able to be used as surrogate endpoints. Whilst this had been accomplished for a small set of classical IHC markers, these were susceptible to technical and interpretive biases and represented only a tiny proportion of potential molecular alterations. Jeselsohn et al.[197] sought to test for procedural bias in routine neoadjuvant comparisons by analysing an extended set of 147 breast cancer-related genes, including the oncotypeDx and PAM50 signatures, in 23 matched and untreated core and excision biopsies. The group's primary focus was on validating Ki67 as a tangible end point for neoadjuvant trials, testing for variation in expression independent of treatment, as well as comparing prognostic signature genes to determine any relative benefit of CB or EB in pre-treatment prognostic assignment. Minimal variation was observed between biopsy time points, with samples routinely clustering with their biopsy pairs for both the oncotypeDx and PAM50 signatures, demonstrating that, whatever variation did occur in lieu of treatment it was less than the variation between non-paired samples. Similar results were obtained for Ki67, with the group commenting that

> "The majority of the genes, including MKI67 (Ki67) did not differ significantly between the core and excisional biopsies without therapeutic intervention, [...] changes seen in Ki67 levels after a brief exposure to a treatment, are likely due to the exposure and not because of differences in sampling"

Whilst genes previously linked with therapeutic response were unaffected by sampling method, Jeselsohn and colleagues also highlighted a set of 14 immune-related transcripts that were significantly altered between core and excision biopsies. Of particular note was *CD68*, a tumour-associated macrophage (TAM) marker. TAMs have previously been associated with poor prognosis, increased grade and decreased disease free survival[198,199], and TAM recruitment in response to CB was highlighted as a real risk, both in confounding the assessment of immune modulating drugs and, perhaps more worryingly, in potentially encouraging factors associated with poor survival.[i]

[i] Faithless lyrists I have read, they'd amputate the rose to know the rose (Gregory Corso)

With one eye on the hype once associated with microarrays, the known and highly convoluted impact of heterogeneity and the huge potential of the neoadjuvant window as a research tool, it remained pertinent to fully explore its dynamics, particularly in the absence of treatment, where a rigorous control study was lacking.

1.11 Hypothesis and aims

Cancers, in particular breast cancers, are not all equal. Classification, as well as prediction of a likely disease course, at the time of diagnosis has allowed more appropriate treatment on a patient-by-patient basis. However, though sub-categorisation has proven useful, it has as yet failed to capture all clinically relevant phenotypes, and patients routinely fall into a category that they do not belong to. This translates to under- and over-treatment. Recent appreciation of just how heterogeneous breast cancers are, as well as improved technologies and approaches have fuelled ever finer characterisation. Two of these — high-throughput molecular assays and neoadjuvant investigation — promise to achieve this characterisation both more reliably and more rapidly. The application of these methods are wide, in particular they allow researchers to integrate clinical and molecular features over time, in particular in addressing the effects (or lack of effect) of treatment. To an extent, the limit to which questions can be asked is the data available and, if the right patient cohort are available, it's possible to investigate many difficult to model problems, not least the acquired resistance to a previously successful treatment.

Therefore the aims of this thesis are three-fold:

1. To model breast cancer temporal heterogeneity, with an emphasis on its relevance to previous and future studies of drug response and biomarker study.
2. To demonstrate the neoadjuvant window's potential in modelling hard to study phenotypes, with a specific focus on acquired endocrine therapy resistance and breast cancer dormancy.
3. To characterise these phenotypes at the molecular level with the intent to derive therapeutic targets as well as predictive and prognostic biomarkers.

2 | Methods

The methods detailed here relate only to work directly performed by myself or in direct collaboration with others. Some work pertinent to the results, though not performed by myself, is therefore not included unless strictly necessary. The thesis as a whole makes use of both novel and previously published datasets, and only those datasets produced *de novo* for the purposes of this thesis are described in detail. All data originates from microarray technologies and quantifies global gene expression, though a number of different platforms were used. All analysis was conducted in R http://www.r-project.org using software packages available via CRAN http://cran.r-project.org/ and Bioconductor http://www.bioconductor.org/.

2.1 General methods and concepts

2.1.1 Microarrays

At the basic level, microarrays are a bed of messenger RNA (mRNA) capture probes. For each expressed protein-coding mRNA, a sequence-specific complimentary oligonucleotide probe is present. As a specific mRNA is introduced, it is able to bind to its complimentary probe and hybridise. Conversely, if a gene is not present (i.e. it is unexpressed) its capture probe remains empty. By fluorescent labelling of mRNA prior to hybridisation, mRNAs can be quantified in response to laser excitation, where individual mRNA quantity is proportional to fluorescent emission.

Whilst microarrays offer the opportunity to greatly increase the number of parallel assays performed, their bulk nature are liable to introduce a number of systematic technical biases and artefacts, collectively referred to as noise. Key to the concept of technical noise reduction is the preservation of true biological signal. Quality control is implemented as several steps, beginning at the point of scanning. A microarray chip typically consists of 6 spots, each composed of the capture probe library and able to perform sample hybridisation in parallel. By comparing spot signal intensity and consistency to background level, proprietary scanning software is able to minimise intra-chip noise. Affymetrix microarray chips in particular include a mismatch probe pairing for each capture probe, designed

to quantify background hybridisation. By varying the capture probe sequence by a single nucleotide, mismatch probes are able to effective represent non-specific mRNA binding, relative to the capture probe.

2.1.2 Fresh-frozen vs. formalin fixed and paraffin embedded material

A key facet of sample collection is storage. Historically, clinical tumour samples have been fixed with formalin before being embedded in paraffin wax, whilst more recently snap freezing in liquid nitrogen has become increasingly common[200]. Fuelling this transition is the molecular degradation observed in FFPE stored samples, particularly in the case of a notoriously delicate molecule such as RNA[200]. This is clearly evident in Figures 2.1.3 and 2.1.4, wherein FF and FFPE samples are integrated via batch correction. Whilst intra-material correlations are high, the correlation *between* materials are significantly lower, most problematically in the case of replicate samples biopsied from the same tumour at the same point in time. The process of batch correction for these samples therefore becomes a balance between maximising the number of samples at our disposal and minimising the influence on artificial correction techniques liable to introduce unwanted and unexpected noise. A further consideration is that to combine FF and FFPE batches a researcher must take the intersection of those probes called with sufficient certainty following microarray hybridisation. As probe detection is known to be particularly poor in FFPE material, FFPE samples can drive a reduction in the available features that could otherwise be preserved if FF samples alone were used in the downstream analysis. Finally, at a practical level, it's increasingly likely that FFPE samples would have been collected significantly earlier than FF samples, and that the corresponding differences in clinical practice and patient lifestyle for FFPE and FF samples may be sizeable and represent a massive confounding factor. Ultimately, the importance of maximising dataset size has been chosen to be of greater value for the analysis performed in the subsequent chapters and FF and FFPE samples were integrated where possible.

2.1.3 Batch correction

Technical artefacts result from variability in all stages of sample extraction, preparation and hybridisation. These artefacts, or "noise", lack any biological meaning, are spurious, cloud interpretation, and perpetuate throughout all stages of downstream analysis if left uncorrected. Key to ensuring a robust analysis is maximising the signal-to-noise ratio, wherein true biological effects are accentuated and technical noise is reduced. Correction for batch effects differ from standard normalisation procedures, described in section 2.1.1, in their per-feature approach, eschewing the notion that all features will be affected equally. ComBat, used throughout this thesis, is an empirical Bayes-based method and, as such, deals

well with small sample sizes[201] and has been demonstrated to outperform other methods including batch mean centring and distance weighted discrimination[74].

2.1.4 Dimensionality reduction

Whilst visualisation of two or three variables can be achieved in as many geometric dimensions, representing hundreds or thousands (e.g. an entire transcriptome) presents a challenge to interpretability. Dimensionality reduction techniques, such as principle component analysis (PCA), allow simplification of high-dimensional multivariate data to greatly reduced linear combinations of variables explaining the greatest degree of variation — the eponymous principle components — often depicted in a human-friendly 2-dimensions that capture the inherent structure of the data. Whilst depiction of raw high-dimensional data is problematic, mathematically solving which linear combination of features are responsible for the greatest amount of variation is trivial. Assuming that these principle components explain a sufficient amount of the total variation within the data, plotting using the 1st and 2nd allows a 2D representation where, in the case of transcriptomic data, samples sharing similar expression profiles will cluster.

Capturing high-dimensional structure for 2D visualisation by PCA therefore allows visual appreciation of global cluster relationships but its linear nature can perform poorly in representing local relationships in the raw data. Alternatively, t-SNE (**t**-distributed **s**tochastic **n**eighbor **e**mbedding) attempts to maintain these local relationships by minimising the divergence between pairwise similarity scores of the raw input and their corresponding low-dimensionality counterparts. In essence, t-SNE iteratively rearranges a random low-dimensional data representation until it's pairwise similarities most closely corresponds to those of the original input's. t-SNE is particularly useful in preserving local high-dimensional structure when displayed in a low-dimensional space, though it may fail to accurately — and quantitatively — represent global differences, with the distance between sample clusters not necessarily depicting their quantitative (dis)similarities.

2.1.5 Clustering methods

Hierarchical clustering reorders variables (e.g. genes, samples) into clusters based on similarities, to reveal shared patterns of expression. By calculating all pairwise similarities the two most correlated features/variables are aggregated as a single cluster and the process repeated using this cluster as a novel pseudofeature in the next round of calculations. Throughout this thesis, "similarity" is calculated as the Pearson correlation coefficient and pseudofeatures are compared using each cluster's least correlated member (complete linkage). Clustering can be represented as a dendrogram, a diverging tree diagram that places variables from the same cluster adjacently. Similarity between individual variables or ag-

gregated clusters are encoded by dendrogram height, with more similar variables being connected by lower horizontal connections. By combining hierarchically clustered sample dendrograms with an expression heatmap, a 2D expression matrix representation, patterns of gene expression can be visually determined and compared.

2.1.6 Differential gene expression analysis (DGEA)

Differential gene expression analysis allows class-specific differences in gene expression to be quantified, generating a measure of statistical significance in conjunction with a magnitude of fold change. Within this thesis, significance analysis of microarrays (SAM) was utilised. SAM non-parametrically computes the strength of association between gene expression and a response variable, such as treatment class or time point, implementing permutation based re-sampling strategies whereby test statistics are compared to a null distribution for random class shuffling.

2.1.7 Pathway analysis

Analysing gene expression patterns at the pathway level allows changes in individual genes to be summarised as a collective effect. This allows the functional significance of multiple genes contributing to a given biological pathway/effect to be combined, simplifies long genelists to smaller and more manageable collections of pathways, and helps explain the most likely biological context in the case of pleiotropy. In essence, for a given genelist, related pathways are summarised and compared as an expected likelihood of being associated with a given pathway by random chance, dependent on the size of the genelist, the pathway and background universe being compared. Within this thesis, statistical enrichment was performed by hypergeometric testing, applied in identifying Reactome pathways in which a gene list of differentially expressed genes were over-represented[202].

2.1.8 Survival analysis

Modelling an association between gene expression and time to a given outcome/event requires a collection of functions able to cope with the realities of clinical trials, that is — patient drop-out and the non-normality of time-to-event data. Survival analysis can be broken up into a collection of inter-related non-parametric, semi-parametric and parametric functions that build upon the survival and hazard functions to determine the association of a variable, such as the expression of a gene of gene signature, and time to outcome[203]. The survival function

$$S(t) = P(T > t) \tag{2.1}$$

where T = case-specific survival, t = time, focusses on the probability of failure for a given survival duration, i.e. the probability of an event occurring later than time `t`. The hazard function

$$h(t) = \lim_{\Delta t \to 0} Pr(t \leq T < t + \Delta t | T \geq t) \Big/ \Delta t \tag{2.2}$$

determines the instantaneous potential (numerator) per unit time (denominator) for event occurrence, given a survival up until at least `t`. A common analogy for the hazard function is a speedometer: at a given instant, a speed of 60mph is simply the potential of travelling 60 miles within a single hour — the true distance covered will vary as the speed increases or decreases around 60mph. The survival and hazard functions are related in that we can reciprocally calculate one from the other. Put simply, as the hazard function increases, the survivor function decreases.

The Kaplan-Meier formula — alternative known as the product-limit formula — estimates the survival function by calculating the product of all probabilities of surviving past a specified time[204]. Whilst the survival function is defined by a smooth continually declining curve, the Kaplan-Meier estimator is represented by a series of decreasing steps, effectively comparing survival time against the proportion of samples (e.g. patients) yet to suffer an event. Critically, the method is able to deal with right censored data, where current outcome information is unavailable and only a record of survival up until a given past time with no indication of current outcome status.

By plotting the Kaplan-Meier estimate, it is possible to represent a cohort's declining survival as a whole or in competition between sub-populations, such as treatment and placebo trial arms. The hazard ratio allows the relative survival of these groups to be compared, in effect deriving a single value for the difference between the two arms.[i] Complimentary to the hazard ratio, survival differences can equally be compared by logrank test calculation, whose formulation is based on the null hypothesis that two groups share identical survival functions[205]. Any differences in survival functions are thereby calculated as measure of significance, implicating the stratifying factor as being associated with outcome

2.2 Data generation and chapter-specific analyses

2.2.1 Chapter 3 : No-intervening treatment

The no-intervening treatment dataset was composed of 74 sequential paired diagnostic core-needle biopsies (CB) and surgical excision biopsies (EB). Thirty-seven patients were originally identified at the Western General Hospital, Edinburgh, between 2003 and 2011,

[i] i.e. group `b` suffer an event at twice the rate of group `a`.

with a histologically confirmed primary invasive breast cancer who had received no pre-operative/neoadjuvant treatment at any time for medical or personal reasons and were entirely treatment-naïve. All patients gave informed consent to be included in the study, which was approved by the Lothian Regional Ethics Committee (2001/8/80 and 2001/8/81) and all experiments were performed in accordance with relevant guidelines and regulations. A summary and complete clinicopathological characteristics of the patients diagnostic core biopsies are given in Appendix 1: Tables.

Core-needle biopsies were taken at diagnosis in all patients using a 14-gauge automated needle device. For each biopsy, multiple cores were taken per tumour and combined as individual samples. Surgical excision biopsies, also physically taken as core-needle biopsies, were collected between 13 and 53 days later (mean interval = 27.5 days). Samples were snap frozen in liquid nitrogen and stored at −80 °C before homogenisation and RNA extraction using the RNeasy Mini Kit with RNAse Free DNAse treatment (Qiagen). Only samples with ≥50% cellularity and ≥60% tumour tissue, determined by haematoxylin and eosin (H&E) staining were included in the study. RNA quantity and quality was assured using a Nanodrop 2000c spectrophotometer (Thermo Scientific). RNA was reverse transcribed and amplified using the WT-Ovation FFPE System Version 2 (NuGEN), purified using the Qiaquick PCR purification Kit (NuGEN), biotinylated using the IL Encore Biotin Module (NuGEN), and purified using the minElute Reaction Cleanup Kit (Qiagen). At each step RNA/cDNA quantity and quality was assured by repeat assessment with the Nanodrop 2000c. Labelled cDNA was hybridized to Illumina Human HT-12 version 3 and version 4 whole-genome expression bead arrays according to the standard protocol for NuGEN labelled samples. Data was extracted using Illumina GenomeStudio software. Illumina probe profiles were pre-processed using the lumi package, consisting of log2 transformation, quantile normalisation. Probes mapping to multiple official gene symbols were summarised by mean expression. Probe expression information was extracted and detected probes were standardised, by passing a detection threshold ($p \leq 0.05$) in ≥3 samples.

The NIT dataset was originally hybridised as 3 distinct processing batches and it was therefore necessary to determine the presence of any technical batch effects and, if present, to correct. Due to the presence of replicate and universal human reference RNA (UHRR)[i] samples between batches, it was possible to accurately infer batch effects by t-SNE visualisation, where it would be anticipated that replicate and UHRR samples should cluster tightly if batch effects were absent. Visualisation using the 500 highest-variance genes within the 3 batches, revealed batches A and B to be highly similar to one another, whilst batch C was noticeably distinct (Figure 2.2.1A). UHRR and replicate samples shared between batches A

[i] UHRR was developed by the MAQC project as a highly reliable, consistent and commercially available reference sample[68].

or B were effectively identical. Though batch C didn't contain UHRR samples, it was clear that replicate NIT samples were distinct from their batch A/B counterparts. In considering overcorrection as necessary to avoid as under-correction, it was determined to consider the 3 processing batches as only 2 batches for correction. Following correction, the three batches formed a single NIT sample cluster (Figure 2.2.1B). Importantly, the biologically distinct UHRR samples are found separated as a second small cluster.

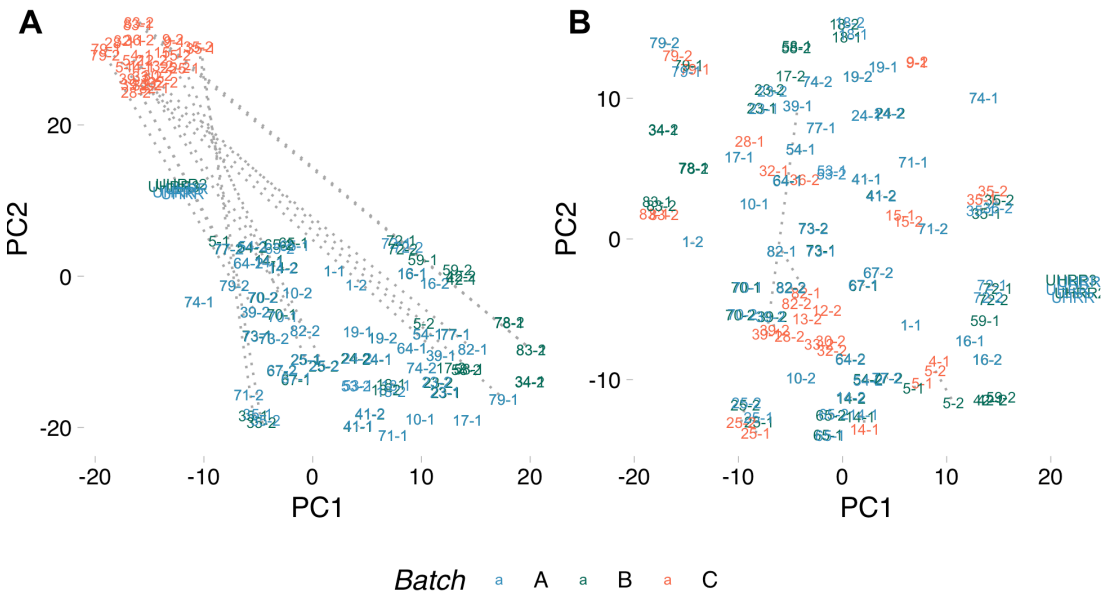

Figure 2.2.1: NIT batch effect removal. Pre-corrected raw data (A) reveals two major clusters, batches A+B (blue, green) and batch C (orange). Replicate samples shared between batches A and B demonstrate high similarity with near perfect t-SNE overlap. In contrast, batch C remains more distal, particularly for replicate samples, where dashed lines display replicate connections. Post-ComBat correction (B) effectively reduces inter-batch differences, noticeably minimising replicate sample distance and maximising NIT and UHRR cluster independence.

Sophisticated batch correction methods, such as ComBat, promise to retain biological signal at the expense of technical noise[201]. Therefore, it is possible to quantify the impact of batch correction as the removal of signal noise and the protection of biological differences. These two metrics were considered as overall correlation, and the combination of intra-batch paired and unpaired correlations respectively. In essence, a 'good' correction would increase the overall correlation (representing removal of technical noise) but not affect intra-batch correlations (representing the preservation of biological signal). Conscious of the distorting impact of replicate samples in batch correction, they were necessary to be removed prior to ComBat correction, accomplished using a hierarchical selection scheme. Given the obvious batch differences, wherever possible batch C samples would be discarded and only selected if they were unique, to minimise the need for correction. For

selection between batches A and B, biopsy correlations would be calculated on a per-patient basis. The replicate demonstrating the highest mean correlation would be retained.

Following replicate sample removal and ComBat correction, it was evident that overall sample correlations significantly improved (p = 1.1x10^{-6}), though not at the expense of intra-batch correlations, indicating the removal of technical noise but not biological signal (Figure 2.2.2).

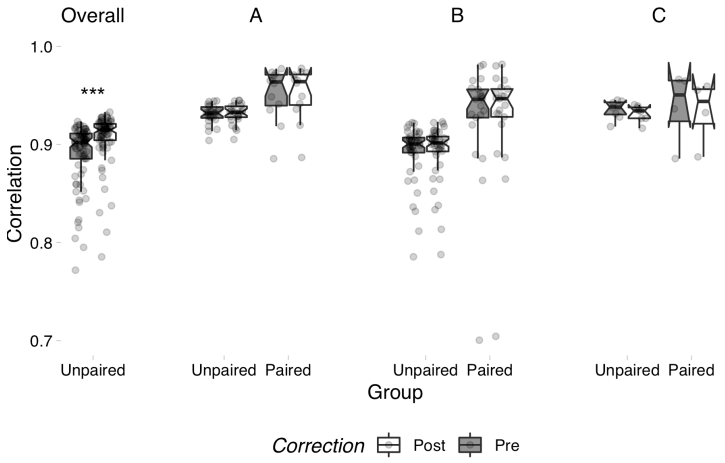

Figure 2.2.2 Batch effect removal quality. Following batch correction by ComBat, overall sample correlation (Overall, representative of technical batch differences) significantly increased (p = 1.1x10^{-6}), though intra-batch paired and unpaired sample correlations (A-C, representative of biological differences) remained unaffected. White = uncorrected, grey = ComBat corrected

All significances were calculated by a two-tailed Wilcoxon rank sum test and corrected for multiple testing (Benjamini-Hochberg correction), unless otherwise stated. Pairwise differential gene expression (DGE) was calculated using significance analysis of microarrays (SAM). Hierarchical clustering was performed on pairwise fold change expression values using a complete linkage method. SAM analysis of treated data was performed using a hundred 18-sample permutations, to fairly match sample size between subsets, with the intersecting significant genes being taken as a mean average. The raw and processed data from this study can be accessed from NCBI GEO under the accession GSE76728. For the letrozole treated comparison dataset, as well as all other datasets used in the meta-analysis, details and sources can be found in Table 3.1.

2.2.2 Chapter 4 : Extended endocrine therapy

The extended endocrine therapy dataset was composed of two independent cohorts composed of multiple sequential biopsies: discovery (n = 177) and validation (n = 53). All patients gave informed consent to be included in the study, which was approved by the local regional ethics committee (07/S1103/26, August 2007). The patients included in the study forwent surgery due to either medical or personal reasons, though all were treated for at least 4-months with letrozole (Femara, 2.5mg; Novartis Pharma AG, Basel, Switzerland).

Over this period, sequential tumour biopsies were taken using a 14-gauge core-needle prior to a final surgical biopsy of the excised tumour. Following biopsying, samples were snap-frozen in liquid nitrogen (discovery and validation cohorts) or formalin fixed and paraffin embedded (validation cohort). Only tumours histologically confirmed to contain $\geq$ 50% cellularity and $\geq$ 60% tumour tissue were included for microarray hybridisation. In parallel, multiple sequential tumour volume measurements were performed by ultrasound sonography (USS). Tumour tissue was homogenised using a membrane disruptor and phase separated. RNA extraction was performed using a column-based purification allowing labelling and microarray hybridisation.

Though sample collection and RNA extraction were performed in Edinburgh, the extended endocrine dataset was profiled as two major batches, one each in Edinburgh and Georgetown, Washington DC. The Edinburgh discovery dataset was labelled and hybridised using Illumina protocols for HumanHT-12 v4 BeadChips and normalisation performed similarly to the no-intervening treatment dataset, consisting of quantile normalisation, log2 transformation, summarisation and probe filtering ($p \leq 0.05$ in ≥ 3 samples). The final discovery cohort was produced in combination with data from two previous studies also generated at Edinburgh: GSE59515[102] (14 patients) and GSE55374[103] (9 patients).

The Georgetown validation dataset, was hybridised using Affymetrix HG-U133Plus 2.0 microarrays. Probes were filtered using MAS5 presence/absence calls before fRMA and loess normalisation.

As each batch was hybridised using different microarray platforms, the Edinburgh and Georgetown datasets would be analysed separately, for discovery and validation purposes, thereby avoiding the complications of inter-platform batch correction. However, each dataset alone remained likely to contain technical batch effects requiring correction, and batch correction would therefore be performed on a per-dataset basis. For the Edinburgh, dataset correction was required to offset technical bias introduced by simple processing batch differences, being a composite of three highly similar small cohorts. The Georgetown dataset was composed of a mixture of fresh frozen (FF) and formalin-fixed paraffin-embedded (FFPE) sample preservations and would require correction on this basis (Figure 2.2.3A). Both for brevity, and that the Edinburgh data was originally pre-processed and corrected by Dr. Cigdem Selli, correction is only detailed for the Georgetown dataset here.

As in Chapter 3, post-correction increases in sample size would be balanced against biological integrity, with care shown to preserve the original intra-batch relationships. Once again, correction quality would be assessed as maximising overall inter-batch correlation (indicating the removal of technical noise) and minimising changes in pairwise correlations between samples from the same batch or patient (indicating the preservation of biological

relationships).

Following ComBat correction for sample material type, the noticeable FF-FFPE batch effect was removed, with samples aggregating by patient post-correction (Figure 2.2.3B). This effect was most noticeable among replicate samples shared between FF and FFPE batches (in all but 2 cases), demonstrating overwhelming preservation of biological signal post-correction. Of the two failures to aggregate by patient following correction (samples 416-1 and 188-1), replicate samples were available that did cluster effectively by patient, clustering with their corresponding paired samples (Figure 2.2.3C). Therefore these outliers were chosen to be excluded in preference of their well-clustered replicates.

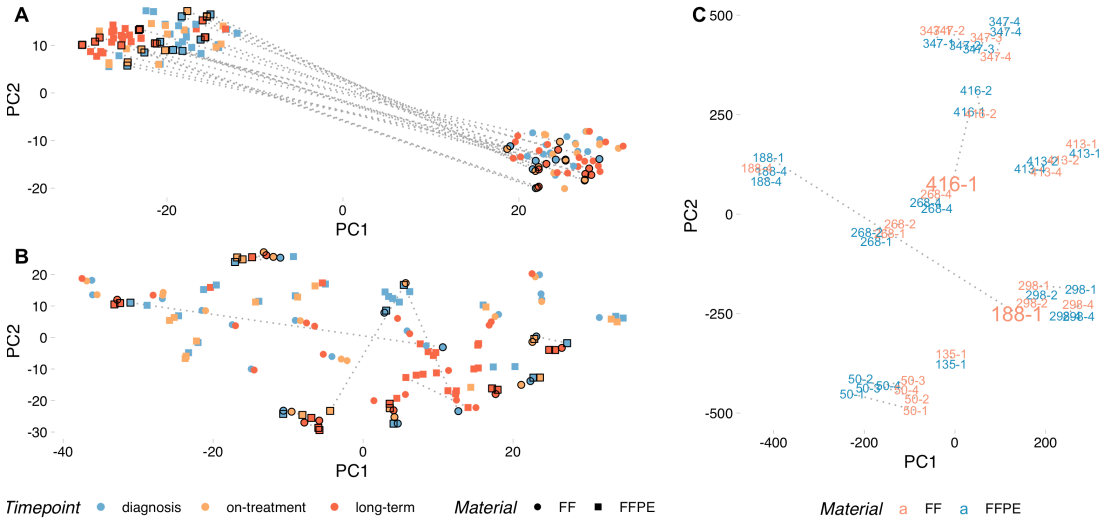

Figure 2.2.3: Validation dataset batch correction. (A) t-SNE dimensionality reduction and visualisation reveals a clear batch effect between fresh frozen (FF, circles) and formalin-fixed paraffin-embedded (FFPE, squares) sample preservations, most notably for replicate samples (black outlines, grey connections). (B) Following ComBat correction, samples cluster irrespective of preservation type, and replicate and paired samples form local clusters of similarity. (C) When replicate samples alone are visualised samples overwhelming cluster by replicate and patient. Only a single replicate pair (416-1 & 188-1, emphasised) exhibit distinct dissimilarity and were removed from subsequent analysis.

ComBat correction was quantified by considering overall and replicate inter-batch correlations as well as intra-batch correlations (Figure 2.2.4). It was clearly evident that as inter-batch correlation improved, intra-batch correlations remained largely unaffected, with post-correction correlations displaying no significant difference (p = .8, one-way ANOVA). Together these results supported increasing sample size by integrating FF and FFPE materials. Aside from the 2 outliers described above, study samples were then selected by removing FFPE replicates where FF equivalents were available, with long-term FFPE storage associated with greater RNA degradation, before re-correction using only those samples included in the downstream analysis.

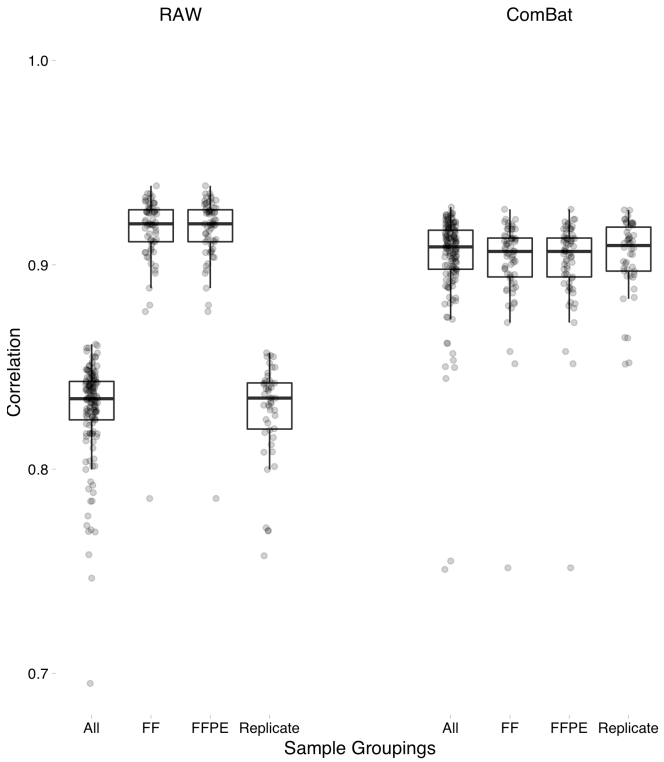

Figure 2.2.4 Batch correction quantification. ComBat seeks to remove technical batch effects whilst maintaining biological signal. To validate its effectiveness, overall and replicate correlations (i.e. samples shared between batches) are used to measure the effects of technical noise; intra-FF and intra-FFPE correlations (i.e. samples from each batch individually) are used to measure the preservation of biological signal. Overall and replicate correlations dramatically increase following ComBat correction (right panel), whilst FF and FFPE correlations remain stable. No significant difference in correlations are evident following ComBat correction ($p = .81$ one-way ANOVA).

Following batch correction, discovery (Edinburgh) and validation (Georgetown) datasets were composed of 177 and 53 samples, detailing multiple biopsies from 62 and 20 unique patients, respectively. A proportion of these samples were shared between datasets and would eventually be partitioned for predictive modelling, though all possible Edinburgh samples were utilised for the initial exploratory analysis.

2.2.3 Chapter 5 : *survivALL*

Chapter 5 used only publicly available microarray data to describe and benchmark the *survivALL* methodology and approach. Relevant information can be found under the GEO accessions: GSE9195[206], GSE6532[207], GSE12093[208], GSE2990[65] & GSE17705[209], or at cBioPortal[43]. The five GEO datasets were identified as part of a systematic literature review in which 44 Affymetrix hybridised breast cancer datasets were found and their phenotypic information harmonised. Importantly, each of ER status, treatment, grade and survival information — amongst other factors — were available following harmonisation. This allowed the totality of all data collected to be queried by these factors, revealing the 5 datasets suitable for the *survivALL* study, defined by 100% ER-positivity, 100% tamoxifen treatment and available survival and grade information. All GEO datasets were downloaded as normalised matrices, though each were then scaled by subtraction of the feature-wise mean and division by the feature-wise standard deviation, to better allow inter-dataset comparison. More so-

phisticated batch correction was purposefully not implemented to impart an additional 'noisiness' handicap to the comparisons.

The METABRIC dataset was selected as it represents the largest single breast cancer gene expression dataset, totalling 1,971 samples with complete disease-specific survival information. It was originally assembled as part of the IC10 subtypinging study[43], as described in section 1.4. Importantly and unusually, METABRIC is split into two equally sized and composition-matched subsets, allowing for independent discovery (n = 980) and validation (n = 991). METABRIC data was downloaded as normalised gene expression data, before scaling as described for the GEO datasets above.

3 | Tumour sampling method can significantly influence gene expression profiles derived from neoadjuvant window studies, *or* **No-intervening treatment (NIT)**

3.1 Context

Breast cancer clinical studies within the adjuvant setting are subject to limitations in speed, patient drop-out, and *in situ* characterisation of response. Alternative treatment prior to surgery, within the neoadjuvant setting, can avoid these limitations, allowing fast determination of preliminary treatment response within a ~2-week to 3-month window. Due to the neoadjuvant window's pre-surgical nature, the *in situ* tumour can be interrogated during the course of treatment, and these clinical measurements related to clinical response. At a simple level, these may include measurements of tumour volume, but the opportunity for multi-sample in-depth molecular characterisation is equally possible[210]. In acquiring multiple biopsies per patient over the course of a study, and in determining those genes most indicative of a response to therapy, the potential exists to develop predictive and prognostic signatures, to better aid patient stratification and, by extension, correct assignment of treatment and the avoidance of over treatment[102]. However, derived signatures are subject to biological as well as practical heterogeneity, that is — the particular cohort assembled and the design, including length, of the study.

Due to the short course of the neoadjuvant window, robust endpoints such as recurrent, metastatic, and disease-specific death incidences remain unknown and alternative measures of efficacy are required. Whilst physical measurements such as tumour volume allow for determination of a pathological complete response (pCR),[i] these still may fail to

[i] "pCR is defined as the absence of residual invasive and in situ cancer on haematoxylin and eosin evaluation

occur within the short neoadjuvant 'window-of-opportunity' study time frame. Therefore, proposals to use reduction in tumour proliferation as measured by Ki67 expression,[i] sought to formalise molecular changes in response to treatment as viable primary endpoints[175].

Ki67 reduction as primary endpoint has been used in a number of neoadjuvant window studies[213,214], comparing pre-treatment diagnostic core-needle biopsies (CB) and peri-treatment surgical excision biopsies (EB), and demonstrated to be associated with both recurrence-free and overall survivals[175,176]. In essence, the studies discussed here perform an initial diagnostic biopsy, administer treatment, and perform a subsequent biopsy at the time of surgery. Though biopsying of the post-operative tumour can take place by a number of methods, for the purposes of this study an excision biopsy was performed by a taking core-needle biopsy of the surgically excised tumour specimen. Within this model, variation occurring between biopsies is attributed to the treatment given, under an assumption of no change in the absence of treatment. Concordance between IHC measured biomarkers (e.g. ER, PR, HER2) has been studied extensively[180], and found to be comparable between biopsies, but these studies routinely lack in scope, either in terms of the markers assayed or the number of samples included. The largest study prior to the publication of *Tumour sampling method can significantly influence gene expression profiles derived from neoadjuvant window studies*[1], analysed 147 breast cancer-related genes, derived from ER-related, PAM50 and oncotype Dx genesets, determining concordance between sample pairs from 21 individual patients[197]. High overall concordance was demonstrated, notably for the PAM50 and oncotype Dx prognostic signatures, however at the same time significant differential expression was demonstrated for a subset of "immune-related" genes. Worryingly, this implicated the core-needle biopsy in activating this immune response, itself associated with poor prognostic factors in breast cancer.

The utility of the neoadjuvant window was therefore threatened by, at best, potential unreliability and at worst detrimental effects to patient well-being. It was therefore necessary to perform a large-scale and high-throughput study to robustly determine neoadjuvant window study paired biopsy sample variation in the absence of treatment. To this end a cohort of 37 patient-derived core-needle and excision biopsy pairs were profiled using Illumina BeadChip microarray technology. Beyond simple characterisation of sample pair variation, it remained critical to determine whether diagnostic prediction by notable prognostic and predictive signatures and biomarkers were liable to differ depending on when during the course of clinical management a biopsy was taken.

---

of the complete resected breast specimen and all sampled regional lymph nodes following the completion of neoadjuvant systemic therapy".[211]

[i] Expressed at all stages of the cell cycle (excluding G_0), Ki67 has been robustly correlated with breast cancer prognosis as an alternative clinical end-point.[212]

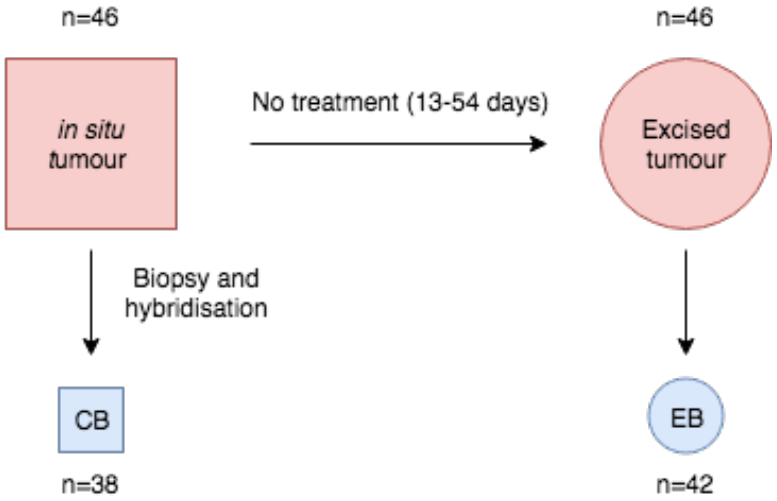

Figure 3.1.1: No-intervening treatment study design. Forty-six sequential tumour pairs (red) were biopsied at both diagnosis (squares) and immediately following surgery (circles). No treatment was given before or during this time frame, covering a span of 13-54 days. Successful biopsying and microarray hybridisation (blue) was apparent for 38 diagnostic and 42 post-surgical samples, of which 37 overlapped and were included in the study.

3.2 Results

3.2.1 Consistent treatment-independent transcriptomic variation within the neoadjuvant window

A perhaps clumsy yet useful null hypothesis[i] was the assumption that under conditions of no-intervening treatment, sequential paired tumour biopsies would display a minimum of transcriptomic variation. Indeed, unsupervised hierarchical clustering revealed a reasonable degree of sample pair co-clustering at the first level of the dendrogram (25/37 pairs, Figure 3.2.1A upper) and pairwise, intra-tumour, correlations were significantly higher than inter-tumour unpaired correlations (p = 7x10^{-11}, Figure 3.2.1B). It was therefore readily apparent that, though samples derived from the same patient were highly similar, they remained non-identical and below the expected dissimilarity as a result of technical interference and sample processing, with microarray-derived correlations below $r \approx 0.97$ likely due to underlying tumour heterogeneity[216].

Perhaps more substantive, a 50-gene differentially expressed geneset (DEGs) was derived comparing paired diagnostic core and surgical excision biopsies (SAM, fdr = 5%). This geneset was enriched for MAPK signalling (*DUSP1*, *JUN*, *NR4A1*, and *FOS*), cancer-specific (*COX-2*, *PGE2*, *JUN*, and *FOS*), apoptotic-induction (*FOS* and *JUN*), and genomic reformatting (brain) ischaemic (*EGR1* and *JUN*) pathways. As expected, based on these genes, samples clustered irrespective of their patient and represented a consistent biological alteration

[i]Though an assumption that was nonetheless implicit in numerous neoadjuvant window studies.

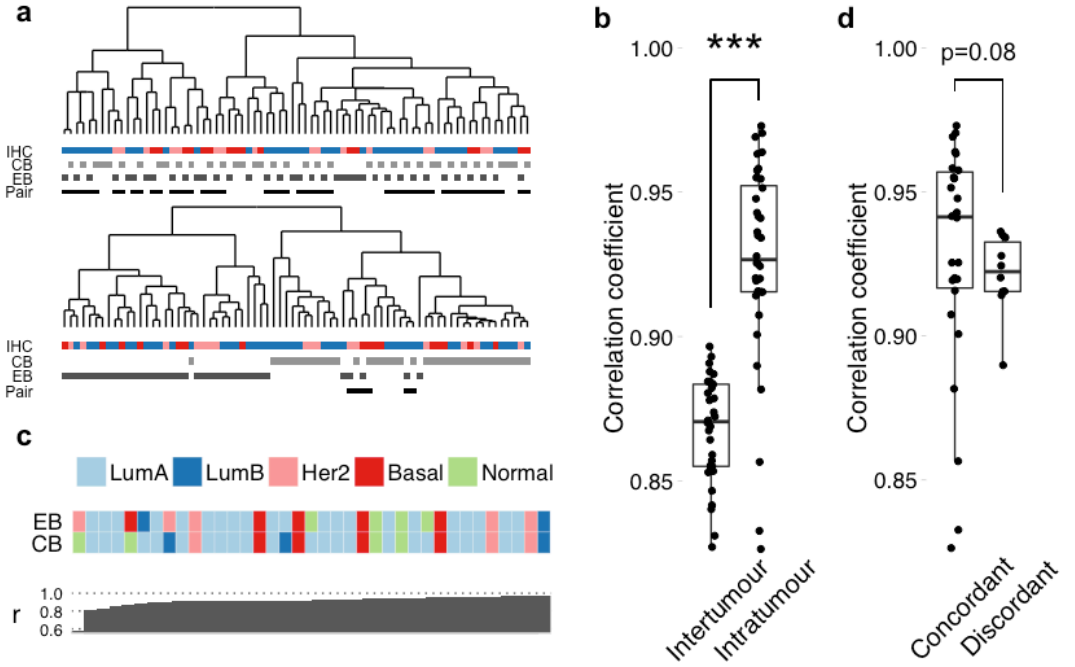

Figure 3.2.1: Treatment-independent and pairwise variation. (a) Hierarchical clustering of 37 patient-matched diagnostic core and excision biopsy samples using the 500 highest variance genes (upper) and a 50-gene DGE signature between core and excision biopsies (lower). Bars represent IHC status (ER+/HER2− = Blue; ER+/HER2+ = Pink; ER-/HER2− = Red) or biopsy method. Lower-most bar indicates where sample pairs co-aggregation. Two-thirds (25/37) of the pairs cluster at the first level of the upper dendrogram, whereas pairwise association is lost in 31/37 cases for the lower. (b) There remains significantly stronger correlation between biopsy pairs (intra-tumour) than between different tumours (mean inter-tumour). ***$p < 0.001$. (c) Discordance in molecular subtype assignment between core and excision biopsies. Patients are ranked left to right by pairwise correlation. Colours represent single sample predictor (SSP) subtypes (Luminal A = Dark blue; Luminal B = Light Blue; Her2 = Pink; Basal = Red; Normal = Green). (d) Sample pairs were called discordant when Biopsy A ≠ Biopsy B for at least 4/5 classifiers. Comparison of concordant vs. discordant pairwise correlations then revealed an inverse relationship between correlation and discordancy.

over the period of no-intervening treatment (Figure 3.2.1A lower). Variation in lieu of treatment was substantial enough to cause discordance in subtype assignment between biopsies. Application of 5 subtyping schemes (PAM50, Sorlie[217], Hu[34], Desmedt[218], and Wirapati[79]) classified patients as either concordant (pairwise subtype agreement for >3 classifiers) or discordant (≤3 agreements, Figure 3.2.1C, PAM50 shown). Per-class pairwise correlations revealed a trend of decreased correlation for discordant patients (p = .08, Wilcoxon signed-rank test).

3.2.2 Pairwise variation in lieu of treatment may be associated with either time or biopsy methodology

Having determined that significant and consistent pairwise variation was occurring in lieu of treatment, it was pertinent to determine the contributing factors. A number of hypotheses appeared likely, that (i) greater variation would be apparent at shorter time intervals if variation was induced in response to diagnostic biopsying; (ii) longer time intervals would display greater divergence if it reflected tumour evolution; (iii) variation was independent of tumour biology. Modelling pairwise correlations as a function of both time interval and IHC subtype displayed no significant association however, largely ruling out a molecular response to diagnostic biopsying as well as tumour evolution (Figure 3.2.2A). Here we would expect an association with time in the case of the former and associations with both time and subtype for the latter, with ER-/HER2- subtypes known to be more mutationally active[219]. Together, these results suggested that the root of the observed variation may not be explained simply as a result of unchecked tumour biology and, following this logic through, it stood to reason that the variation observed without treatment would likely be equally apparent in a similarly processed treated dataset. By applying the 50-gene NIT signature to a letrozole treated dataset, analogous in all of laboratory, microarray platform and structure (multiple paired biopsies over time) it was obvious that a subset of the treated samples displayed a similar pattern of expression to those of the NIT dataset (Figure 3.2.2B). The treated samples that most exhibited the NIT expression pattern were those with the longest treatment, implicating time between biopsies as the causative factor. However, as in the NIT dataset, these longer-term treated biopsies understandably happened to be those that were collected at the time of surgery, a fundamentally different biopsy method.

The definitive root cause of NIT signature expression — either time or biopsy method — was difficult to dissect. For the letrozole treated dataset, upregulation of the NIT geneset was apparent only after 3-months, for both on-treatment core-needle biopsies as well as surgical excision biopsies, albeit to a lesser degree for CB samples (Figure 3.2.2D). However, global geneset enrichment analysis (GSEA) analysis was able to demonstrate that the NIT signature could significantly define the differences between untreated and treated data only when sampling method differed (NIT vs. 2wCB, $p = 0$; NIT vs. 3mCB, $p = .02$; NIT vs. 3mEB, $p = .25$). Similarly, only in the instance of excision biopsy was DGEA able to reproduce ≥30% of the NIT geneset. More so, out of 3,955 differentially expressed treated-EB genes, seven of those common to NIT differential expression were amongst the top 15 in terms of fold-change magnitude. It therefore remained difficult to disentangle whether time between biopsies or biopsy method was the causative factor in NIT geneset expression and in all likelihood a combination of both may have been apparent.

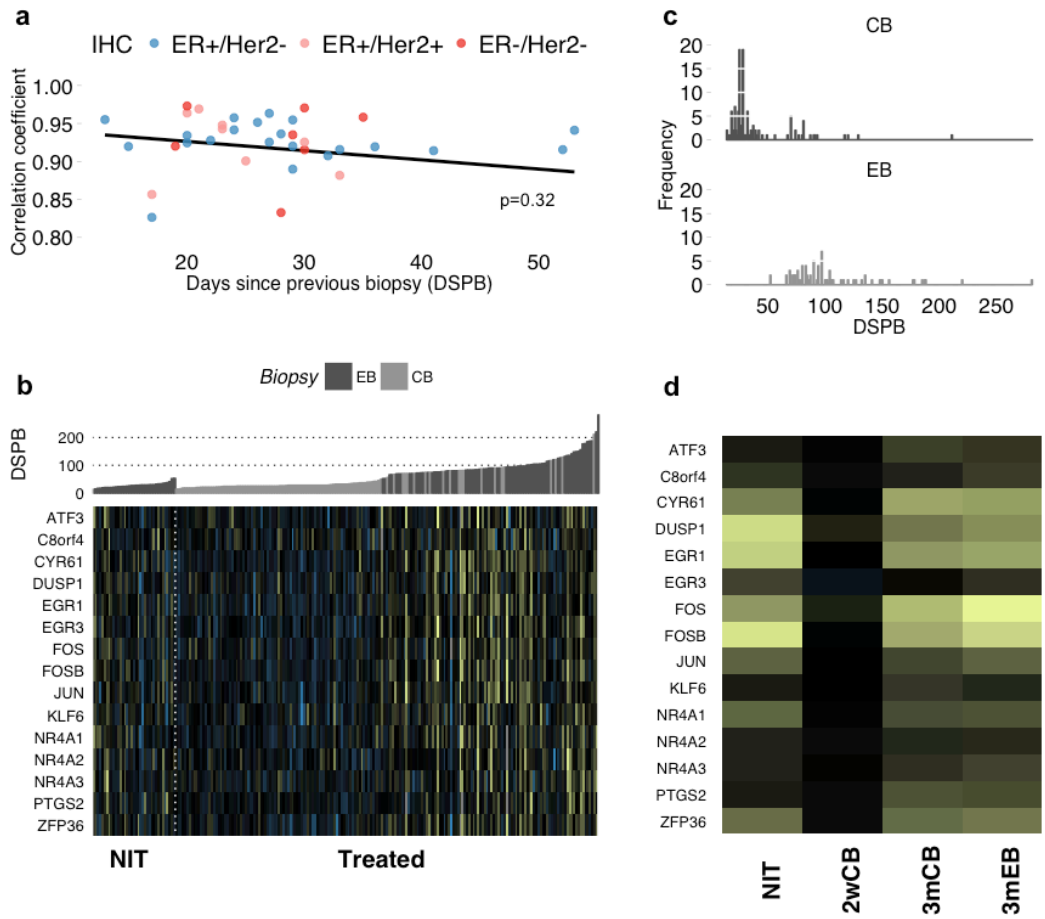

Figure 3.2.2: Causative factors influencing treatment-independent variation. (a) Pairwise correlations between biopsy pairs are not explained as a function of either time between biopsies ($p = .32$) or IHC status ($p = .43$). ER+/HER2− = Blue; ER+/HER2+ = Pink; ER-/HER2− = Red. (b) NIT signature fold-change between CB and subsequent patient-matched EB biopsies for both NIT and letrozole treated cohorts (upregulation = yellow; downregulation = blue). Samples are ordered by increasing time between biopsies and reveal a pattern associated with either extraction method - CB (grey) or EB (dark grey) - or time. (c) Frequency distribution of biopsy time intervals. (d) To investigate the effects of biopsy method and/or time on gene expression, the letrozole treated data was split into three subsets – 2-week CB (2wCB), 3-month CB (3mCB) and 3-month EB (3mEB) and mean fold changes calculated. Both 3-month subsets closely resemble NIT differential expression.

3.2.3 Multiple patient-matched datasets also demonstrate changes in NIT early growth response genes

Irrespective of the underlying reason for pairwise variation in the NIT geneset, it remained that changes occurring in lieu of treatment were apparent in treated datasets. To further validate the NIT signature, a panel of four genes most representative of pairwise NIT gene-set variation (*DUSP1*, *EGR1*, *FOS*, *FOSB*) were selected based on differential expression fold-change magnitude and significance, for comparison in six external validation datasets

(Figure 3.2.3 & Table 3.1). Importantly, as well as being highlighted in the NIT geneset, these genes were both well characterised in the literature as early growth response genes and notable for their potentially spurious inclusion in previously published neoadjuvant studies[220–222]. For these genes, significant differential gene expression was consistently observed to a greater degree only when an excision biopsy followed a previous core-needle biopsy in all datasets.

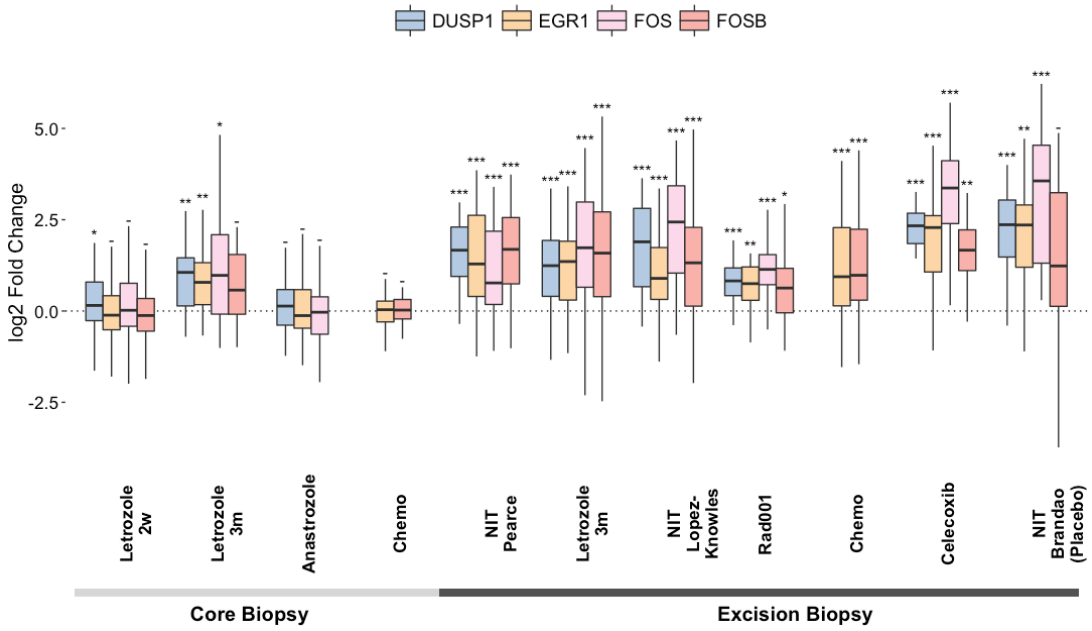

Figure 3.2.3: Pairwise variation for four early growth response NIT signature genes in six validation datasets. These genes potentially represent an association between gene expression and sampling method, with surgically excised samples (EB) showing greater expression fold-change than their core-needle biopsied (CB) counterparts. $*p \leq 0.05$; $**p \leq 0.01$; $***p \leq 0.001$; $-$ = not significant.

		Biopsy Time Interval/days		Biopsy Type			
Treatment	*n*	median	range	1st	On-	Final	Dataset/Reference
None	37	27	13–53	CB	-	EB	Pearce 2016
Letrozole	122	107	13–884	CB	CB	CB/EB	Turnbull 2015
Celecoxib/none	22/15	NA	14–21	CB	-	EB	Brandão 2013
Anastrozole	99	14	14–112	CB	CB	CB	Smith 2007
RAD001	21	14	14	CB	-	EB	Sabine 2010
Chemotherapy	69	NA	NA	CB	CB	EB	Magbanua 2015
None	56	14	14	CB	-	EB	Lopez-Knowles 2016

Table 3.1: Composition of patient-matched neoadjuvant breast tumour datasets used in NIT signature meta-analysis.

3.2.4 Markers of proliferation as surrogates for clinical response to treatment

Given the International Ki67 in Breast Cancer working group's recommendation of on-treatment changes in Ki67 expression as an alternative to pCR as well as longer term outcome measures[212], it was necessary to determine whether *MKI67* gene expression was liable to change in lieu of treatment. However, unfortunately, probe quality control deemed *MKI67* to have not been reliably measured in the NIT dataset. Despite this, it remained possible to compare core-needle and excision biopsies using alternative proliferation markers, such as *PCNA*[223], *MCM2*[224] and *CCNB1*[225] (Figure 3.2.4). Expression of these three markers clearly varied, demonstrating a general decreasing trend between biopsies, though not significantly so (Figure 3.2.4A, Wilcoxon signed rank test p-value range = 0.07-0.35). This tendency for proliferation to decrease was significantly more pronounced (*PCNA* $p = 1.1\text{x}10^{-10}$; *MCM2* $p = 4.0\text{x}10^{-16}$; *CCNB1* $p = 1.9\text{x}10^{-23}$, one-way ANOVA) in the letrozole treated neoadjuvant dataset of similar experimental design. As a positive control, 2 NIT signature genes (*HBB* and *EGR1*) were also compared, with the degree of differential expression remaining similar between datasets (Figure 3.2.4B).

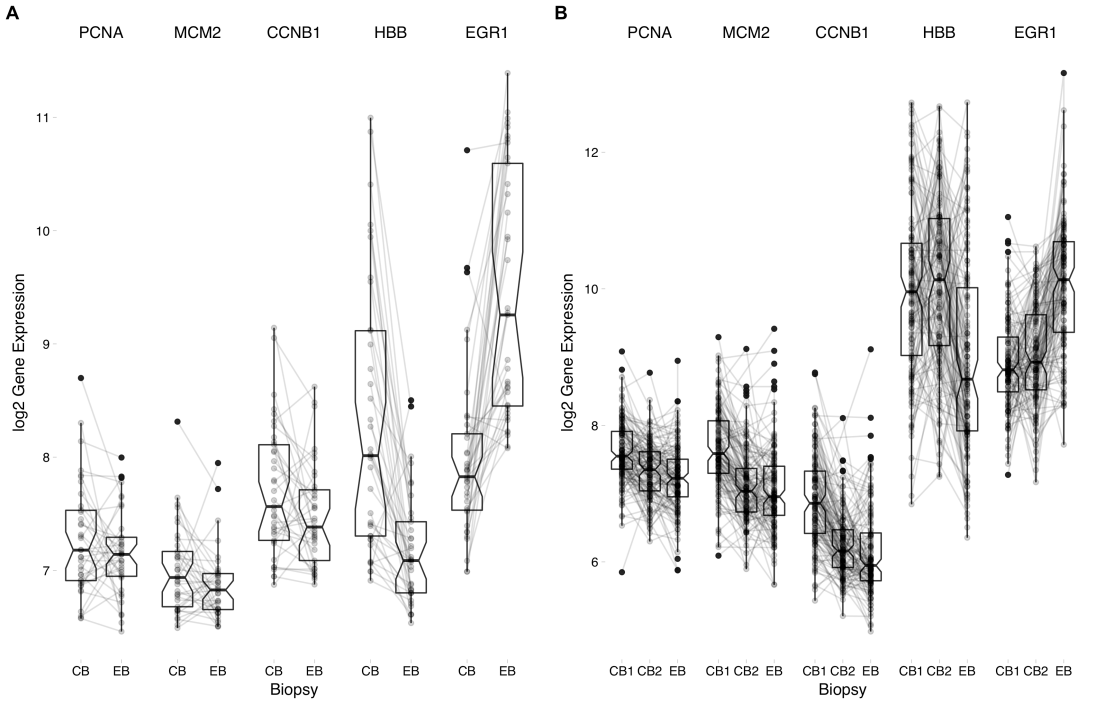

Figure 3.2.4: Reduction in proliferation in lieu of treatment. Three markers of proliferation (*PCNA*, *MCM2* and *CCNB1*) demonstrated non-significantly decreased expression between core-needle and excision biopsies in the NIT dataset (A) (Wilcoxon signed rank test). These relative differences were more pronounced, and independent of biopsy type, in a second treated dataset (B) (*PCNA* $p = 1.1 \times 10^{-10}$; *MCM2* $p = 4.0 \times 10^{-16}$; *CCNB1* $p = 1.9 \times 10^{-23}$, one-way ANOVA). For comparison, two NIT signature genes (*HBB* and *EGR1*) were similarly significantly differentially expressed in both datasets. Lines denote pairwise changes.

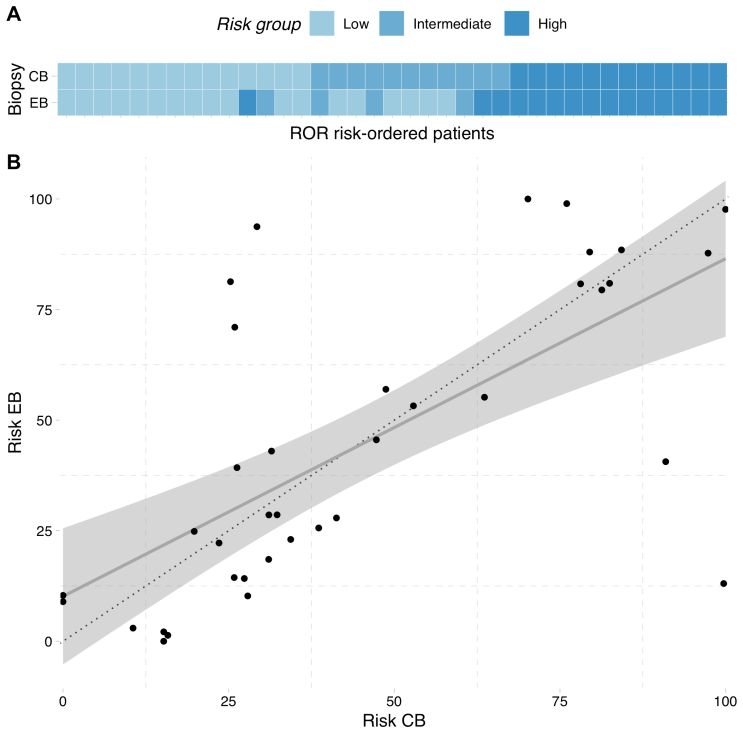

Figure 3.2.5: Prosigna risk of recurrence variation between biopsies. (A) Categorical classification is largely robust between biopsies, with a minority of classifications varying. Where risk group classification does alter, it is almost exclusively interchange between one of either low or high risks and intermediate risk. (B) Changes in risk classification are mediated by variation in risk scores between biopsies.

3.2.5 Prognostic assignment by commercial gene tests is partially affected by biopsy type

From a clinical perspective it was important to determine whether biopsy time and/or type could affect prognostic assignment by clinically available gene signatures and, by extension, whether either biopsy could be considered more suitable in determining prognosis. Prosigna ROR classification displayed discordance, though only in a minority of patients (29%) and — for all but a single patient — only to or from the intermediate classification (Figure 3.2.5A). Core-needle and excision biopsy risk score correlation revealed a moderate correlation (Figure 3.2.5B, R^2 = .46, p = 4.2x10^{-6}), explaining the variation observed in Figure 3.2.5A.

3.2.6 Immune-related response to core-needle biopsy

Finally, immune activation, specifically recruitment of the poor prognosis-associated *CD68*, as a result of a core-needle biopsy had been previously highlighted[197]. However, the NIT dataset failed to significantly demonstrate this effect for any of the available (7/9) highlighted genes (Figure 3.2.6, pairwise Wilcoxon signed rank test = 0.16-0.83). Furthermore,

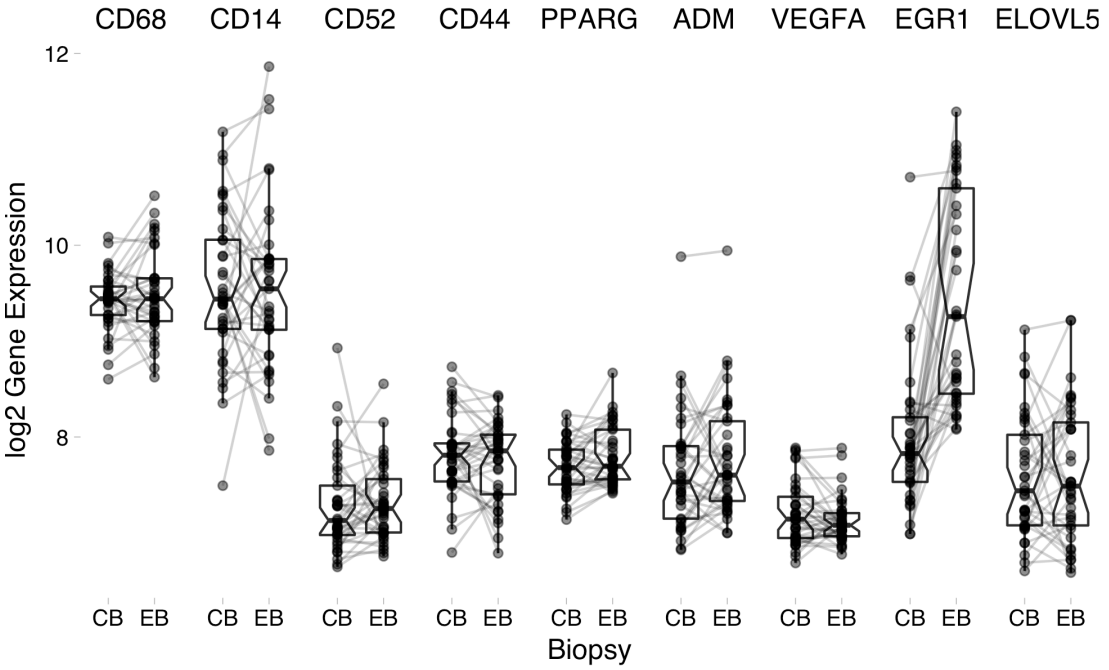

Figure 3.2.6: Lack of immune activation in response to core-needle biopsy. All available immune signature genes (*CD68*, *CD14*, *CD52*, *CD44*, *PPARG*, *ADM*, *VEGFA*) were compared between core-needle and excision biopsies, though none demonstrated a significant change in expression. For comparison, *EGR1*, a NIT signature gene, and *ELOVL5*, a negative control highly expressed in testis, were included. Lines denote pairwise changes.

pathway enrichment analysis demonstrated a lack of any significantly ($p < .05$) enriched pathways, immune or otherwise, for the proposed immune gene set.

3.3 Perspectives

The no-intervening treatment study was an important as well as necessary control analysis to reinforce the applicability of the neoadjuvant window as a tool for expedited clinical studies in breast cancer. The results presented here detail that changes in gene expression do indeed occur between untreated tumour biopsies, within the time frame of the neoadjuvant window, and can potentially introduce a confounding factor into downstream analysis. However, this variation is superseded or "drowned-out" by changes occurring in response to therapy. This informs the validity of previous neoadjuvant studies conducted to date but still cautions against over-interest in novel biomarkers overlapping with the NIT signature geneset, many of which have indeed been highlighted in previous treated studies[220,221,226,227]. It should, however, equally be noted that it remains plausible for the NIT signature genes to have a legitimate role in treatment response in addition to their response to surgical excision.

The proposed causative hypothesis for NIT variation is up-regulation of an early growth

response in response to surgical trauma and warm ischaemia before newly biopsied samples are processed, where the cellular metabolic machinery attempts to mount a survival or apoptotic response before all metabolic activity ceases. Tissue ischaemia may result from exclusion of the vascular supply or simply from handling by the surgeon, scrub nurse, pathologist and tumour bank personnel before the sample is snap-frozen in liquid nitrogen. This time delay is likely (on average) to be significantly longer for surgical excision specimens than core biopsy samples. The effect of ischaemia on gene expression has been described previously and warm ischaemia associated with the surgical extirpation of human tissues has significant effects on gene expression[228]. Morrogh et al.[227] similarly concluded that window trials are influenced by a wound-healing process, comparing a 502 cancer-related gene panel in 8 untreated controls and Dash et al.[229] demonstrated significant changes in the expression of *FOS*, *JUN*, *ATF3* whilst investigating the effect of sample processing time in prostate cancer. Two further studies also highlighted ischaemia as a potential source of molecular variation and, again, key molecules including *FOSB* were highlighted[230,231]. Importantly, the investigators compared excision biopsies placed into RNAlater either immediately post-surgery or following an interval, and again observed early growth and stress response associated genetic expression to increase.

Ideally this ischaemic effect would be confirmed by comparing paired biopsies taken immediately before and after surgery. However, ethical considerations, specifically the previously proposed induction of a poor outcome-associated immune response following a core-needle biopsy, prohibit superfluous biopsies being performed. However, transcriptome-wide DGEA of paired biopsies taken at increasing post-surgical time intervals has subsequently been performed, revealing that post-surgical biopsy processing time does indeed significantly effect gene expression. Importantly, significant differences in many of the NIT signature genes (34%), including *DUSP1*, *FOS*, *FOSB*, *JUN*, *ZFP36*[232] conclusively implicate surgery, surgically-associated sample fixation and X-ray processing times with NIT signature gene induction. Taken together these data support the careful monitoring of ischaemic time for tissues harvested for the purpose of gene profiling.

The clinical impact of neoadjuvant changes in gene expression in lieu of treatment remain somewhat unresolved. The results here suggest that, if excision biopsying induces the NIT response, then core-needle biopsies remain the optimal point at which to consider prognosis by gene signatures and biomarkers. Perhaps the only argument counter to this logic is that intermediate ROR risk classification appears to be depleted for surgical excision biopsies, with classifications being more clinically relevant/actionable as either ‘high’ or ‘low’.

It remains that a major limitation of the study was the necessary exclusion of *MKI67*, among other potentially informative genes, due to probe quality control measures. Reli-

ably demonstrating *MKI67* expression consistency between biopsies would have allowed greater confidence in the use of Ki67 as a surrogate outcome marker. Instead, it was only possible to demonstrate the consistency of three alternative proliferation markers — *PCNA*, *MCM2* and *CCNB1* — none of which were evidenced to significantly alter in a paired analysis, though a trend of decreased expression was apparent. However, when compared with the analogous proliferative decreases expected in a treated study, these changes proved negligible. It has recently been demonstrated that IHC measured Ki67 significantly increases between untreated core-needle and excision biopsied samples, in contrast to sharp decreases in chemotherapy treated patients[233]. It could be argued that this does not invalidate treated studies use of Ki67 as a surrogate outcome measure, as similarly to the results presented here, the no-treatment effects are again clearly superseded by their treated counterparts.

Further methodological considerations include the determination of intra-patient discordancy at the unsupervised level. Here, discordancy was calculated using the combined results of 5 gene expression signatures, each applied in assigning subtype to highlight changes between core and excision biopsies from the same patient. Importantly, the assignments for any given sample were themselves subject to variation between signatures, where one signature may call a sample Luminal A and another call that same sample as HER2-enriched, potentially inflating the perceived differences resulting from sampling methodology (Figure 3.3.1). Of the subtyping models applied, 3 classified samples as 1 of 5 subtypes and the remaining 2 as 1 of 3 subtypes. Understandably the 3-subtype models demonstrated a greater degree of consistency in assigning the same subtype to the same sample, with a large proportion of 5-subtype variation being between Luminal A and Luminal B subtypes that may only differ subtly at the molecular level. For the purposes of the NIT study it may have been more robust to apply only the 3-subtype models in assigning discordancy. Furthermore, though subtype assignment results in a discrete classification it additionally computes the confidence of that same assignment. These confidence values could additionally have been utilised to control for superfluous subtype disagreement.

Beyond the methodology of the study, it remains that, whilst full transcriptome measurement allowed previously unseen characterisation under conditions of no treatment for breast cancer within the neoadjuvant window, clinically it's unlikely for biomarkers to be measured by microarray, with cheaper and quicker alternatives such as qPCR being more practical. It would therefore be prudent to validate the NIT signature in terms of these simpler methods. Furthermore, experimental design mimicking that used by Lopez-Knowles et al.[232], comparing variable post-surgical tissue processing times, could allow a quantitative assay of surgical impact to be devised, measuring changes in NIT signature

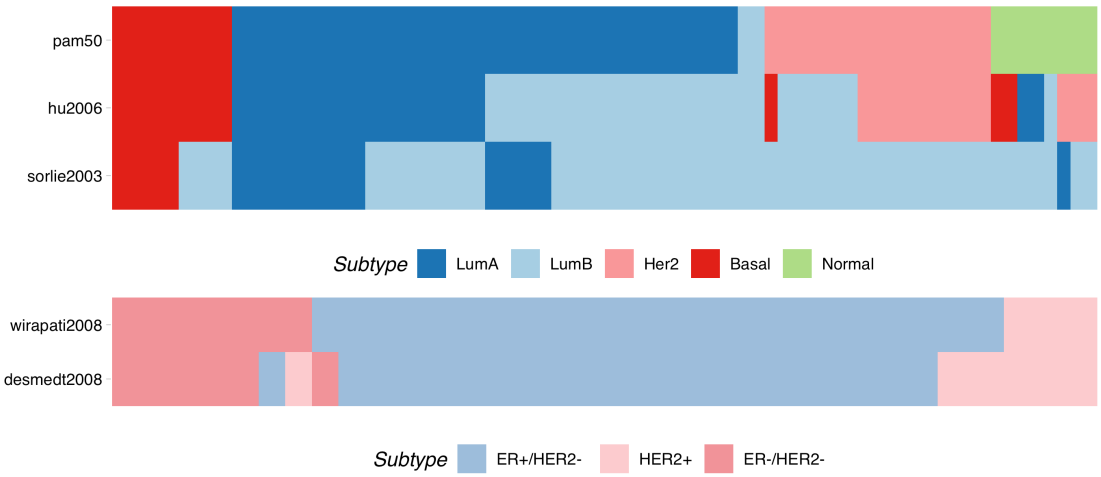

Figure 3.3.1: Aggrement in subtype assignment. Samples, ordered from left to right were subtyped by 5 different methods. Consistency/variation in assignment for the same sample is displayed in the vertical axis. Subtype assignment is observed to vary depending on the method used. Understandably the majority of variation in subtype assignment for the 5 subtype models (upper) occurs between Luminal A and Luminal B classes. This is reinforced by the more stable assignment demonstrated by the 2 subtype models (lower).

genes most representative of inter-biopsy changes independent of treatment. In this way, treated studies could determine the study-specific impact posed by surgical excision.

Ultimately, with the potential for pairwise variation irrespective of treatment, our study raises potential concerns of the suitability of the neoadjuvant window in gene expression profiling studies. Recent results of the ALTTO (adjuvant lapatinib and/or trastuzumab treatment optimisation) clinical trial were, however, found to be consistent with the predicted benefits from the neoALTTO trial[234]. This supported the utility of the neoadjuvant setting as a suitable and important window for evaluating promising new targeted agents, as well as the continued use of patient-matched samples to assess intervention studies for translational research[235]. Nonetheless it remains critical to understand that whilst patient-matched samples reduce variation due to individuals, all possible sources of variation must be considered for an optimal experimental design. For example, in an intervention study to assess dietary changes on normal breast tissue from pre-menopausal women it was considered optimal to schedule the sequential biopsies one menstrual cycle apart, rather than using a fixed window of time[236], as there is clear evidence that menstrual changes in oestrogen levels caused significant changes in gene expression[237]. Underlying tumour heterogeneity is an inevitable variable when comparing any two tumour biopsies, and this study suggests that the method of sample collection should be considered along with treatment, time interval and clinicopathological features as an important potential confounding factor. These considerations are of particular importance if a study's purpose is the development of a prognostic/predictive classifier or identification of a biomarker, with genes

present in our NIT signature excluded from the analysis. Beyond the results highlighted here, the NIT dataset represents a publicly available control, against which new findings within the context of the neoadjuvant window can be compared and benchmarked.

4 | Extended neoadjuvant therapy as a model of breast cancer dormancy and acquired endocrine resistance, *or* **extended endocrine therapy**

4.1 Context

The majority of invasive breast cancers present as oestrogen receptor positive. Therapeutically this represents a double-edged sword, wherein a tumours addiction to oestrogen and ER signalling both drives its proliferation but simultaneously renders it susceptible to targeted therapies. For a majority of cases (50-70%), endocrine therapy is successful, tumour volume decreases within the neoadjuvant setting and post-operative recurrence free survival is extended[104,238,239]. However, for these same initially responsive tumours, the likelihood of recurrence is in excess of ~30%, even for small node-negative tumours[240,241]. This is in contrast to ER-negative disease, which, despite lacking targeted therapies, exhibits a continuous reduction in recurrence risk beyond the first five years post diagnosis (Figure 4.1.1)[153].

The potential for delayed recurrence in ER-positive disease following seemingly successful treatment is currently difficult to understand, as well as to study experimentally. It is conjectured that undetected residual disease is likely to seed future recurrences. This residual disease may represent only a tiny fraction of the original tumour mass, and even a single cell may be able to eventually proliferate and become detectable[164]. As a result, tracking, biopsying and assaying residual and dormant[i] tumour cells is challenging, even outside of the normal challenges of extended follow-up studies such as patient drop-out and cost. Furthermore, prediction of when a tumour cell may emerge from its dormant

[i]That is tumour cells that have, or are yet to, clinically re-emerge from a prolonged disease-free period

state is equally obscure, with residual tumour cells able to remain dormant for a longer duration than would be anticipated by normal cell doubling rates[242]. Studying breast cancer dormancy within the adjuvant setting is therefore fundamentally difficult and, to date, studies have been restricted to cell line and animal models.

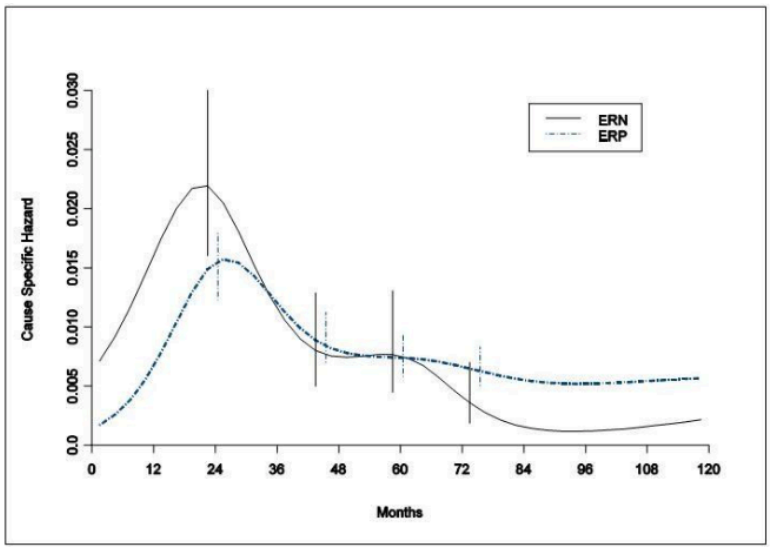

Figure 4.1.1: Distant metastasis hazard estimates for 1615 ER-positive (ERP) and 427 ER-negative (ERN) breast cancers. ER-positive samples display a continued recurrence risk up to at least 10-years (adapted from Demicheli *et al.* 2010).

In contrast, the neoadjuvant window offers an alternative setting to model the dormancy trajectory *in situ*, that is:

(i) treatment → (ii) response → (iii) desensitisation and resistance *or* continued response

Within the framework of the Edinburgh Breast Unit's biobanking activity, a cohort of patients undergoing extended neoadjuvant endocrine therapy was identified. Extended treatment, in some cases reaching over 3.5-years, allowed multiple ultrasound measurements and tumour biopsies to be collected sequentially. With these samples, breast cancer dormancy had the potential to be experimentally interrogated. By confirming an initial response to therapy, determined by a reduction in tumour volume, patients could be categorised as either one of two classes — **dormant** or **desensitised** — based on whether the initially observed volume decrease eventually reversed. This chapter will present an analysis that first validates the appropriateness of extended neoadjuvant therapy as a model for long-term breast cancer dormancy, before characterisation and classification of dormant and desensitised patients based on their transcriptomic profiles.

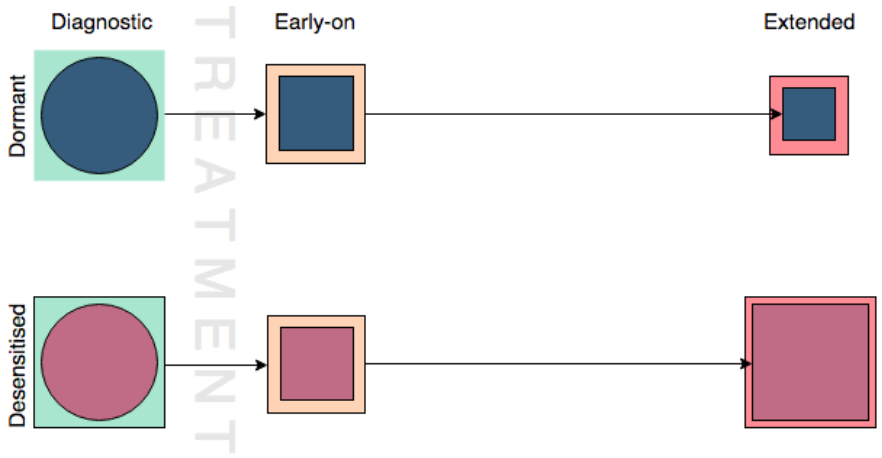

Figure 4.1.2: Dormancy classification and sampling scheme. Three time points (Diagnostic, Early-on and Extended) characterise two tumour classes (Dormant and Desensitised).

4.2 Results

4.2.1 Extended neoadjuvant therapy as an appropriate model for acquired endocrine resistance

Key to studying long-term acquired endocrine therapy resistance was demonstrating the appropriateness of the extended neoadjuvant model. Considerations of treatment duration, change in tumour volume and how these were reflected at the molecular level needed to be carefully assessed. Critical was the demonstration of initial treatment sensitivity followed by either continued sensitivity or eventual desensitivity. This would ultimately allow classification of patients as one of two classes — dormant or desensitised.

Ultrasound sonographic (USS) measurement taken *in situ* over the course of treatment allowed tumour volume to be a read-out of treatment response. An initial decrease in tumour volume of ≥40% within 120 days of treatment was required for inclusion in the study, and a subsequent continuing decline or steady volume classified patients as dormant. Conversely, re-growth of at least 5% classified patients as desensitised (Figure 4.2.1A). Classification by USS was complicated, however, due to a proportion of patients having their final USS measurements taken at least one month prior to their final tumour biopsy. This allowed the possibility of treatment desensitivity occurring within this window to be missed. For these patients (n = 46), transcriptomic measurement of 3 markers of proliferation (*MKI67*, *PCNA*, and *MCM2*) were used to help guide dormant/desensitised classification (Figure 4.2.1B). Average expression was compared across the three treatment time points, that is — pre-treatment (**Diagnostic**); treatment within the common neoadjuvant study time frame (≤120 days, **Early-on**); and treatment beyond this (>120 days **Extended**) (4.1.2). Only where gene expression classification greatly differed from USS measurement was classification recon-

sidered, with USS otherwise being considered before gene expression measurements. Following classification, the Edinburgh dataset consisted of 40 dormant and 22 desensitised patients.

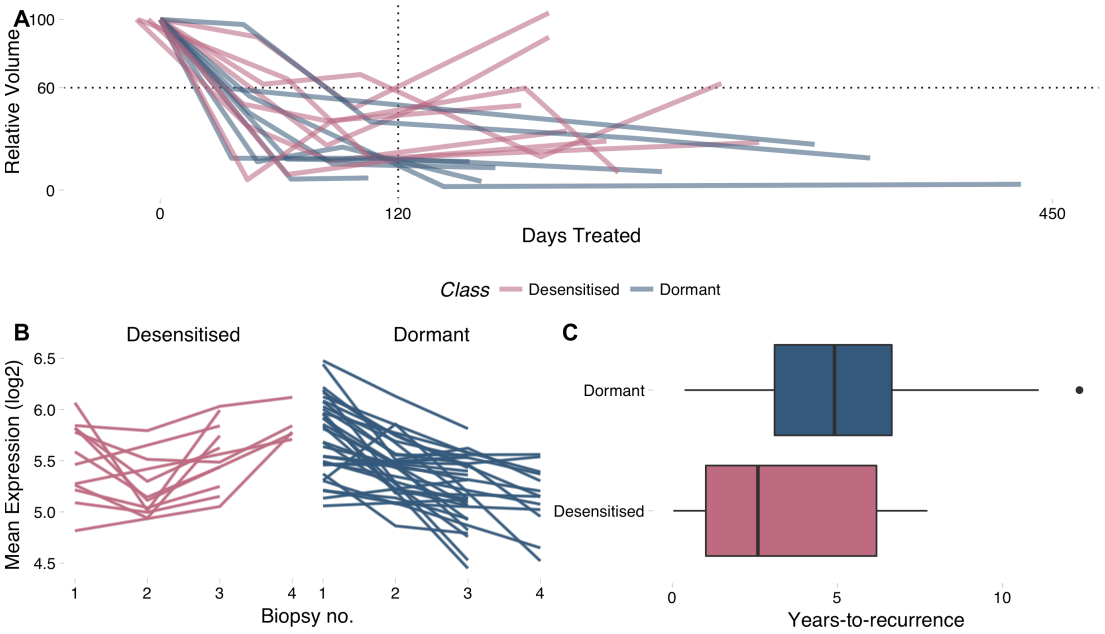

Figure 4.2.1: Treatment response and class assignment. (A) Sequential USS measurements taken over the course of the study allowed patients to be assigned as either dormant or desensitised based on the following criteria: a ≥40% decrease in tumour volume followed by either a subsequent ≥5% increase (desensitised) or continued reduction/stability (dormant). The x-axis is limited to 450 days for visualisation. (B) For patients whose last USS measurement was ≥1-month prior to their final biopsy, proliferation is used as a surrogate for changes in tumour volume. (C) Time-to-recurrence is evident to be significantly shorter for desensitised patients ($p = .01$, one-sided Wilcoxon rank sum test).

Classifications were benchmarked using both clinical and transcriptomic variables. Firstly, time-to-recurrence was determined to be significantly earlier in desensitised patients (Figure 4.2.1C; $p = .01$, one-sided Wilcoxon rank sum test). Then, to validate dormant/desensitised classifications were reflected at the transcriptomic level, two assumptions were tested, that (i) early-on treatment samples would demonstrate less between-class diversity, under the selective pressure of treatment and (ii) dormant and desensitised patients would display variable long-term prognoses. Despite no significant differences evident in clinico-pathological features between classes (Table 6.5), the comparative enrichment for 50 'biological hallmark' genesets[243] between dormant and desensitised patients demonstrated significant variation in gene expression at the diagnostic time point. This variation in between-class enrichments was greatly reduced at early-on treatment time points, demonstrating an initial shared response to treatment. However, by the extended treatment time point, renewed diversity was apparent mirroring the expected variation in treatment trajectory outlined above (Figure 4.2.2A). Of the geneset

evident to differ between classes, early and late oestrogen responses are differentially enriched at both diagnostic and extended time points, but not early-on treatment. Similarly, oncogene-related hallmarks (MYC targets, KRAS signalling, angiogenesis, and DNA repair amongst others) were all differentially enriched at diagnostic and extended time points.

In parallel, PAM50 subtyping revealed a general transition of subtypes only if a patient had been classified as dormant (p = 2.25x10$ˆ{-6}, Fisher's exact test). Anecdotally, this trend appeared to be towards better prognosis subtypes (e.g. Luminal B > Luminal A; Luminal A > Normal) by the extended time point. Desensitised patients on the contrary displayed no evidence of transition between subtypes, largely remaining as the predominant class: Luminal B (p = 1; Figure 4.2.2B & Table 4.1).

		Diagnostic	Early-on	Extended
Dormant				
	LumA	18	19	27
	LumB	23	9	5
	HER2	0	0	0
	Basal	0	0	0
	Normal	0	0	10
Desensitised				
	LumA	7	7	7
	LumB	11	9	11
	HER2	1	1	1
	Basal	0	0	1
	Normal	0	0	0

Table 4.1: Subtype variation over extended endocrine treatment. Dormant and desensitised classes exhibit varying progression under treatment, here illustrated by subtype heterogeneity. As treatment continues, dormant patient subtypes and associated prognosis begin to change ($p = 2.25\text{x}10^{-6}$, Fisher's exact test), transitioning from a majority Luminal B subtypes to a majority of Luminal A/Normal subtypes. Desensitised patients on the other hand display almost no change in subtype proportions ($p = 1$, Fisher's exact test), typified as Luminal B tumours remaining Luminal B.

These results confirmed the *a priori* assumptions regarding the dormant and desensitised phenotypes. Classification was therefore considered successful and the extended neoadjuvant model as a whole appropriate for further investigation.

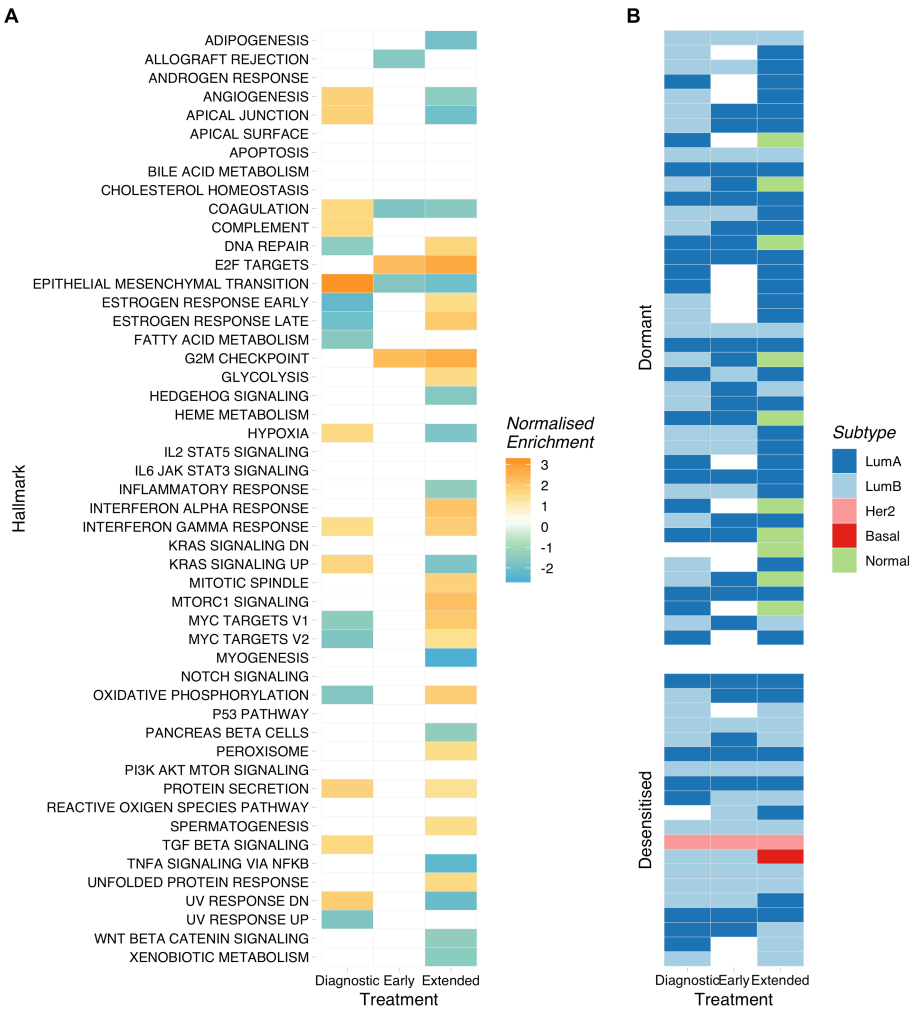

Figure 4.2.2: Molecular validation of extended endocrine class assignment. (A) Normalised-enrichment scores (colour) demonstrate relative enrichment of hallmark geneset between classes. Though patients initially differ for a number of hallmarks at diagnosis, early treatment greatly reduces between-class variation. After extended treatment patients again exhibit increasing variability. Scores are relative to dormant patients. (B) Class-wise PAM50 subtyping reveals a trend of improving prognosis for dormant patients and stable/worsening prognosis for desensitised patients. (C) Though the distribution of subtypes are similar between classes at diagnosis, a clear shift from Luminal B → Luminal A → Normal subtypes is apparent for dormant patients as treatment progresses. In contrast desensitised patients transition from Luminal A to Luminal B or even Basal subtypes.

4.2.2 Class differences are most evident after extended treatment

Transcriptomic class differences were initially compared using unsupervised feature selection and clustering. PCA analysis using the 500 highest-variance genes across all samples, revealed little suggestion of an obvious dormant/desensitised stratification, instead samples clustered based on treatment duration, with treated samples effectively progressively diverging from their diagnostic pairs in the 1st principal component (PC1; Figure

4.2.3A). However, there was the suggestion that this effect was not shared between classes, with only extended treatment dormant samples significantly separating from their their untreated diagnostic pairs based on PC1 (p = 7.9x10^{-14} *dormant*; p = .1 *desensitised*; one-way ANOVA; figure 4.2.3A lower). In essence the expected molecular effect of treatment observed in dormant patients appeared disrupted for desensitised patients and, indeed, the pairwise correlations between diagnostic and extended treatment samples was significantly lower in dormant patients (p = .004, Figure 4.2.3B), suggesting that long-term variation in gene expression may have potential in predicting class membership. For single time points, unsupervised clustering largely failed in distinctly stratifying classes, though, consistent with Figure 4.2.3A, extended treatment samples displayed some hint of class separation based on gene expression, at least for the right-most sub-cluster of potentially "easier-to-call" dormant patients (4.2.3C). However this translated poorly to earlier time points and may simply reflect that class differences are too slight to be resolved early during treatment, or that they simply do not yet exist.

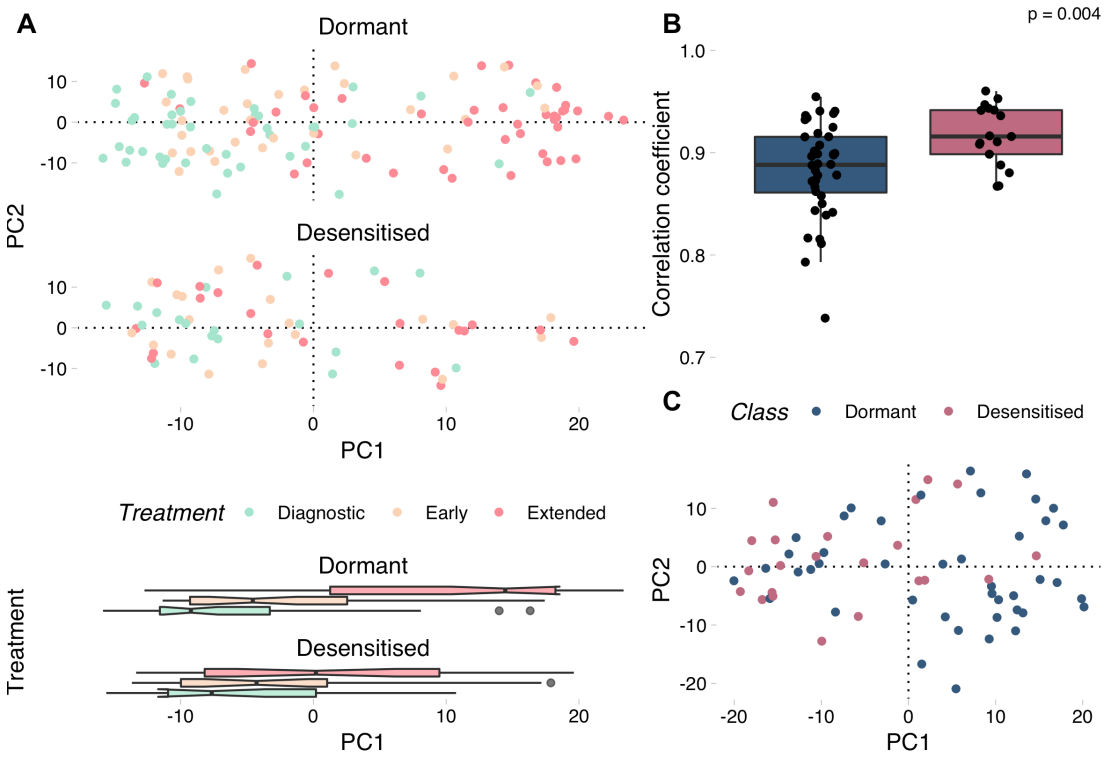

Figure 4.2.3: Class-wise effects of extended treatment. (A) PCA visualisation of the 500 highest-variance genes failed to explicitly stratify classes. Instead samples clustered by time on treatment, most notably in PC1. This representation of treatment effect over time varied by class, with desensitised extended treatment samples co-clustering with earlier treated and diagnostic samples. (B) Quantification of this effect demonstrated a significantly higher pairwise correlation between diagnostic and extended treatment samples for desensitised patients (*p* = .004). (C) Alternatively using the highest-variance genes of extended treatment samples only is suggestive of class stratification based on transcriptomic features.

Per-class differential expression (SAM, fdr < .05) between diagnostic and extended treatment samples revealed further class-wise differences, specifically in the number of genes called as differentially expressed (5466 *dormant* vs. 41 *desensitised*).[i] The greatly reduced number of DEGs in desensitised patients further highlighted the apparent attenuation/reversal of a shared response to treatment (Figure 4.2.3A), and in fact only 4 of the 41 desensitised DEGs (*AATF*, *HBE1*, *MIR193BHG* & *RGS1*) were not shared by dormant patients. Moreover, of these 41 genes, 16 were apparent to overlap with the no-intervening treatment signature geneset, raising the possibility that a sizeable proportion of desensitised DEGs were simply as a result of comparing core and excision biopsies, rather than any specific genetic reprogramming. The intuition here therefore remained that class differences stemmed from the loss of a previously shared response to treatment, rather than an obvious alternative mechanism or program of gene expression. Furthermore, application of dormant paired DEGs to desensitised samples revealed that not only were the dormant changes being attenuated, they were in fact being reversed, an effect that appeared to increase with time-on-treatment (Figure 4.2.4).

Dormant differential expression was typified by the expected pathway involvement derived from previous neoadjuvant endocrine studies, with proliferation, cell cycle and senescence pathways all highlighted. That these changes were absent in desensitised patients again makes intuitive sense and PCA of paired dormant differentially expressed gene fold change was able to detail a reasonably well defined stratification of dormant and desensitised patients (Figure 4.2.5A). Similar stratification was alternatively achieved using DEGs from direct dormant and desensitised extended treatment sample comparisons (Figure 4.2.5B). Pathway analysis of the union of these significantly up- and downregulated genes (n = 234) revealed a notable enrichment for Reactome epigenetic modification pathways at $p < 0.05$ (hypergeometric test), with several of these pathways being similarly enriched in the previous paired analysis, including ‘HDACs deacetylase histones’ and ‘DNA methylation’, again implicating epigenetic modification as a potential determinant of dormancy/desensitisation.

[i] DGEA was additionally performed using 50 random samplings of dormant patients. These samples were each equal in size to the number of desensitised pairs, allowing for any discrepancies in differential expression as a result of differences in sample sizes to be corrected for. The average number of DEGs were 2,121 (range = 909-3,856).

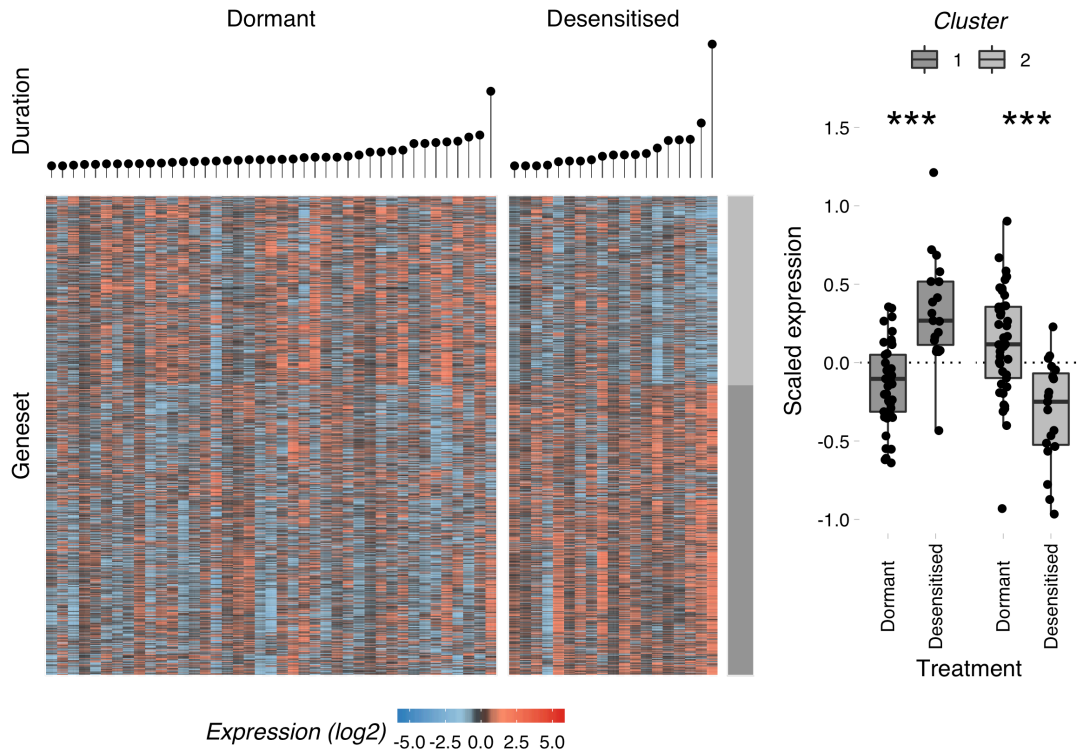

Figure 4.2.4: Treatment induced changes are reversed after extended treatment in desensitised patients. Dormant differentially expressed genes (y-axis) over the course of treatment are not shared by their desensitised counterparts. In fact the opposite effect is revealed, with an increasing magnitude following longer treatment duration. Of the two major gene clusters (light and dark greys), a significant difference in expression fold-change is observed between classes (up: $p = 1\text{x}10^{-6}$; down: $p = 1.9\text{x}10^{-5}$, Wilcoxon rank sum test). Patients (x-axis) are ordered by increasing time-on-treatment. Red = up-regulation, blue = down-regulation. All comparisons are drawn between extended and diagnostic biopsies.

Description	GeneRatio	BgRatio
M Phase	158/2749	325/10281
Chromosome Maintenance	68/2749	110/10281
G2/M Checkpoints	92/2749	170/10281
PRC2 methylates histones and DNA	51/2749	75/10281
Cell Cycle Checkpoints	103/2749	202/10281
DNA methylation	46/2749	66/10281
RNA Polymerase I Promoter Opening	45/2749	64/10281
Signaling by Rho GTPases	187/2749	446/10281
ERCC6 (CSB) and EHMT2 (G9a) positively regulate rRNA expression	50/2749	77/10281
HDACs deacetylate histones	57/2749	93/10281

Table 4.2: Paired differential expression pathway enrichment. Dormant gene expression changes over the course of treatment. All pathways were significant to $p < 1\text{x}10^{-35}$. GeneRatio is the ratio between the number of DEGs in the pathway and the number of mapped DEGs. BgRatio is the ratio between the number of genes in the pathway and the total examined background

Description	GeneRatio	BgRatio
HDACs deacetylate histones	37/142	93/10281
RNA Polymerase I Promoter Opening	31/142	64/10281
DNA methylation	31/142	66/10281
PRC2 methylates histones and DNA	32/142	75/10281
Condensation of Prophase Chromosomes	32/142	75/10281
Activated PKN1 stimulates transcription of AR (androgen receptor) regulated genes KLK3 and KLK3	31/142	68/10281
SIRT1 negatively regulates rRNA Expression	31/142	70/10281
HATs acetylate histones	38/142	143/10281
ERCC6 (CSB) and EHMT2 (G9a) positively regulate rRNA expression	31/142	77/10281
Meiotic recombination	32/142	89/10281

Table 4.3: Class-wise differential expression pathway enrichment. Gene expression changes between dormant and desensitised extended treatment samples. All pathways were significant to $p < 1x10^{-9}$. GeneRatio is the ratio between the number of DEGs in the pathway and the number of mapped DEGs. BgRatio is the ratio between the number of genes in the pathway and the total examined background

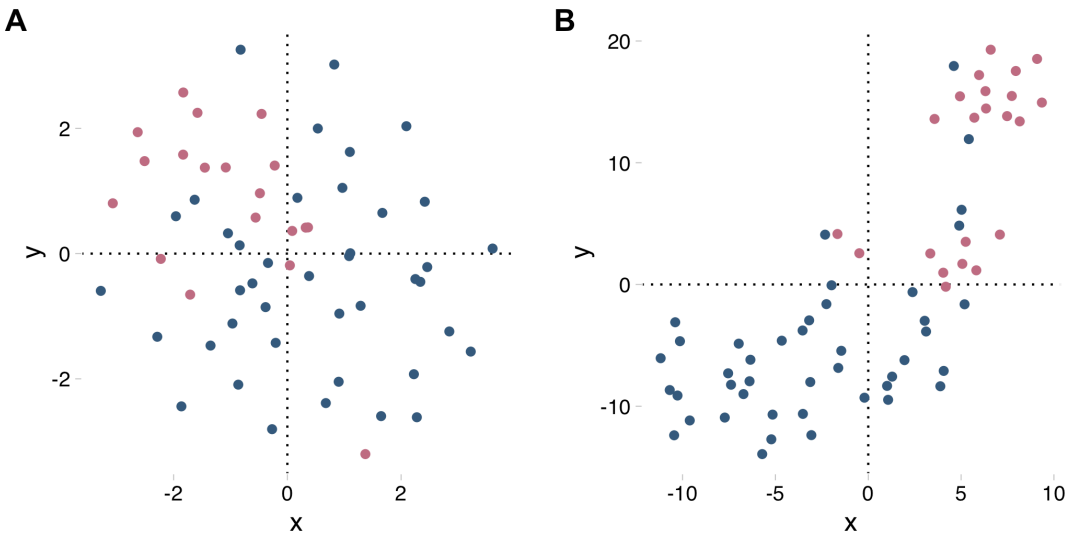

Figure 4.2.5: Class stratification and visualisation by DEGs. t-SNE visualisation of extended treatment samples by dormant paired diagnostic→extended (A) and unpaired between-class (B) DEGs highlight the potential for class stratification and prediction. Blue = dormant; red = desensitised.

4.2.3 Class prediction

Building a class prediction model was therefore attempted using alternative approaches, applying both the paired and unpaired feature subsets individually as well as in combination. A random forest model would be built using the discovery dataset and features refined by variable importance. This model would be trained using a subset of the Edinburgh discovery cohort that was repartitioned to redistribute replicate samples shared by both the discovery and Georgetown validation datasets in an attempt to balance class proportions and retain a sufficiently large validation set. In response to this, differential expression was recalculated using the newly partitioned discovery samples only. Models were trained on a subset of the discovery partition, evaluated using a withheld discovery subset, before application to the validation data (Table 4.4). The aim here was to predict desensitised patients suitable for an alternative treatment regimen.

Data	Dormant	Desensitised	% Dormant
Discovery	24	9	73
Discovery Evaluation	6	3	67
Validation	18	8	69

Table 4.4: Discovery and validation dataset partitions.

On the whole class prediction was unable to predict class (Figure 4.5). The best accuracy (62%) was achieved using fold change values for the differentially expressed geneset derived from dormant patients' paired diagnostic and extended treatment samples, though this fell short of the benchmark accuracy achievable by calling no patients as desensitised (69%). The shortfall in accuracy was driven by poor sensitivity (25%) that characterised the various prediction approaches, where desensitised patients appeared inherently difficult to predict.

		Reference	
		Desensitised	Dormant
Prediction	Desensitised	2	6
	Dormant	4	14

Accuracy:	62%
Sensitivity:	25%
Specificity:	78%
PPV:	33%
NPV:	70%

Table 4.5: Validation confusion matrix. Random forest model prediction fell below a baseline accuracy of 69%, largely driven by low sensitivity in the prediction of desensitised patients. PPV/NPV = Positive/Negative Predictive Value.

4.2.4 Epigenetic modification may mediate desensitisation

Both paired and unpaired differential expression models had highlighted epigenetic-associated genesets, including histone deacetylases, histone acetyl transferases and DNA

methyltransferases, though in reality the overlap between these pathways was high. Nonetheless, it was clear that epigenetic modification and regulation was integral to the molecular variation observed between dormant and desensitised classes. Inspection of the 13 genes shared by the highlighted pathways revealed a significant progressive down-regulation of all 13 genes following treatment for dormant patients (Figure 4.2.6A; one-way ANOVA). This effect was not shared by desensitised patients following extended treatment, with *HIST1H2AK* in fact showing significant and opposite *up-regulation*. For desensitised patients, there was a clear trend for an initial decrease at early-on treatment time points before subsequent re-expression, an overall neutral effect, at least within the study duration. This initial response followed by desensitisation to treatment mirrors the USS and proliferative markers used to assign class membership and offers further evidence to latent acquisition of the desensitised phenotype that is not apparent at diagnostic or early-on treatment time points.

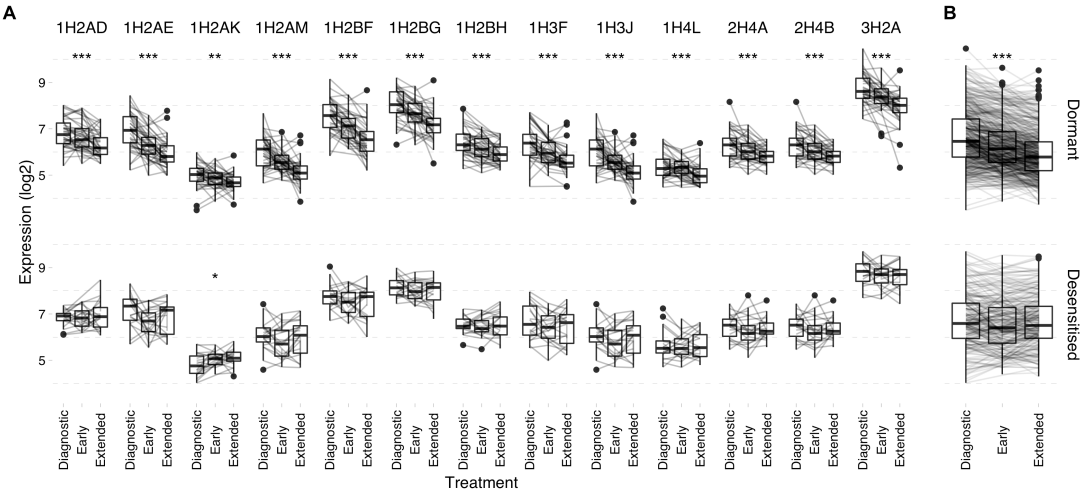

Figure 4.2.6: Class variation in epigenetic-related expression. (A) Thirteen epigenetic-related HIST genes are significantly downregulated over the course of treatment in dormant patients only, summarised in (B). Lines denote pairwise changes. *** = $p < .001$, * = $p < .05$.

The pioneer factors *FOXA1*, *GATA3*, and *PBX1* have each previously been linked with influencing an alternative ER-signalling response via chromatin remodelling and activated ER recruitment and both *FOXA1* and *GATA3* were evident only to significantly decrease in dormant patients, potentially fitting with epigenetic-mediated desensitisation to endocrine therapy[244,245] (Figure 4.2.7).

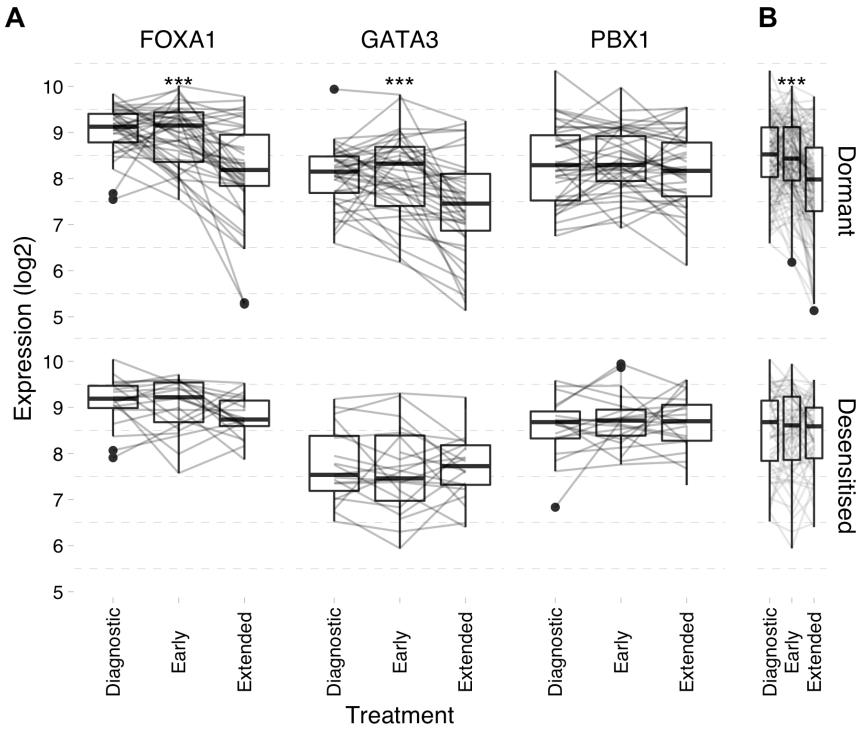

Figure 4.2.7: Class variation in pioneer factor expression. (A) Two of three pioneer factors associated with ER-chromatin interaction are significantly downregulated over the course of treatment in dormant patients only, summarised in (B). Lines denote pairwise changes. *** = $p < .001$.

4.3 Perspectives

The molecular affects of endocrine therapy have previously been demonstrated for intrinsically resistant and responsive patients[102]. Post-treatment alterations in proliferative (↓), metabolic (↓), immune-responsive (↑) and extra-cellular matrix/stromal remodelling (↑) pathways are essentially lost in resistant patients and, moreover, these effects are apparent at 2-weeks on-treatment. Acquired, rather than intrinsic, resistance stands as a different challenge, and patients initially present as ostensibly identical in terms of their sensitivity to endocrine therapy, typified by shared changes in the same pathways listed above. Eventually, however, this sensitivity is lost at some undetermined future point in time. Experimentally this makes acquired resistance particularly difficult to model, reflected in the paucity of patient-derived models of acquired resistance and breast cancer dormancy. Even within the context and normal time frame of the neoadjuvant window, tumours are likely to have been surgically removed prior to desensitisation occurring, making molecular characterisation challenging. However, through normal deviation from routine disease management, some patients forwent surgery for an extended period of time, allowing neoadjuvant treatment to be continually administered and the possibility of desensitisa-

tion to occur. Characterisation of this cohort confirmed that the expected trajectory of an initially shared response followed by either continued sensitivity or desensitivity was confirmed by a number of approaches. At the clinical level, USS demonstrated a decrease in tumour volume followed by stabilisation/subsequent increase and these phenotypic changes were equally reflected at the molecular level. Rational biological differences, as demonstrated using the hallmark geneset[243], were evident between classes. Here it was expected that class differences would be most evident after extended treatment time points and this proved evident, often for logical facets of tumour biology, including cancer-related signalling pathways, such as MYC, KRAS, WNT, and oestrogen signalling pathways (Figures 4.2.2). Proliferative, angiogenic, DNA repair and EMT components were similarly differentially enriched, with angiogenesis in particular suggested to play a major role in emergence from dormancy[165]. Perhaps the most obviously difference between classes was simply xenobiotic metabolism, suggesting that the mechanism of endocrine resistance was related to how drugs were metabolised by the tumour, with dysregulation of *CYP2D6* previously previously being implicated in tamoxifen resistance[115]. Class-wise differential enrichment of the hallmark genesets was clearly extensive after long-term treatment and, working backwards, these changes were almost non-existent early on-treatment, likely as a result of the initially shared response to therapy coercing all tumours to a single similar phenotype before eventual deviation into distinct classes. Though differential enrichment at diagnostic time points tantalisingly suggested the possibility of early class stratification, this never truly manifested in subsequent unsupervised or supervised comparisons.

Intrinsic subtyping revealed similar class variation over the course of treatment, demonstrating a clear trend towards better prognosis phenotype only in the case of dormant patients. It was clear that each of these qualitative measures were evident to progress over the course of treatment in a manner that fit with the dormant/desensitised hypothesis and spoke to the validity of the model as a whole, encouraging further analysis.

Unsupervised characterisation of class variation reflected and reinforced this pattern of molecular change over time. Here treatment was anticipated to inform a majority of the variation observed within the extended endocrine dataset and this was evident, with samples clustering by time on treatment (Figure 4.2.3). Though the effect of treatment could have overshadowed the likely more subtle class differences, in fact variation *in* the response to treatment proved a major discriminating factor, fitting with the overall message that desensitisation manifests as a reversal of the effects of treatment, where long-term treated desensitised samples molecularly revert to an early-on or even pre-treatment state. In illustrating the continued response to treatment it was again clear that an initial response to therapy was shared between classes, before reversion for desensitised patients, with many of these genes hinting at potential mechanisms facilitating desensitisation, including

epigenetic and cell-cycle regulation (Figure 4.2.4 & Table 4.2). It appeared that response to therapy would be more informative than any single biopsy snapshot in time. Moreover, it's clear that desensitisation to endocrine therapy is not simply an attenuation of the anti-tumour effects but that the continued changes seen in dormant are in fact entirely reversed, up-regulation becomes down- and *vice versa*, and this effect becomes more predominant the longer post-diagnosis a biopsy was taken (Figure 4.2.4). This suggests that desensitised extended treatment samples may represent a more aggressive cancer than that at diagnosis.

Whilst this variation was highly promising it failed to generalise and accurately facilitate class prediction. Neither the most promising DEGs derived from dormant treatment-induced changes, nor direct class comparisons following long-term treatment, alone or in combination would translate effectively to the Georgetown validation dataset. Accuracy (59%) was no better than the class proportions (62%), and therefore failed to perform better than simply predicting all samples as dormant. Furthermore, class prediction was performed using extended treatment samples, and performed worse on earlier time-points, limiting any clinical applicability even if successful.

It is perhaps here that the utility of the extended endocrine dataset reaches its limit, as a descriptive, rather than predictive, tool. The uniqueness of the cohort is reflected in its relatively small sample size and a yet smaller still validation dataset. Furthermore, the explicit classification of tumours as either dormant or desensitised is a convenience to allow comparison but likely does not represent to true reality of desensitisation to endocrine therapy. Dormant tumours may simply be yet-to-become desensitised tumours and two truly independent classes may not exist, with initially responsive patients likely to suffer a recurrence up to at least a decade post-surgery. Therefore, though two explicit classes have been considered in this analysis, the designation of dormant and desensitised could equally be considered a single class: initially responsive patients who *have had* and who *will have* a recurrence. As a result, uncovering clear divisions amongst the dormant and desensitised classes would always be challenging, more akin to comparing *yes* and *not-yet* than *yes* and *no*. Classification, here measured by USS would therefore always stand as a snapshot, subject to the effects of temporal heterogeneity, despite being taken as several measurements over time, and a far greater neoadjuvant treatment time span would be necessary to truly compare long-term dormant patients with their desensitised counterparts.

This, in turn, assumes that the use of USS can accurately determine the relatively subtle changes in tumour volume that underpin the dormancy study. Though USS has been commonly used to measure *in situ* tumours, its resolution in finely assessing changes in volume is unclear and the test/re-test accuracy is high though imperfect[170]. It appears that alternative approaches such as CT scan may offer greater confidence in determining tumour vol-

ume, though this comes at the cost of efficiency and ease of use, both for the clinician and the patient. Given that desensitisation was classified using an increase in volume of 5% underlies the precariousness of classification, where to an extent it has been trusted that the measurements are reliable. Despite this, the measurements themselves were performed in a highly controlled manner, within a single centre and by the same clinician, at the very least ruling out the opportunity for inter-observer disagreement. Moreover, though a 5% increase may appear small, this proportion is relative to the original tumour volume and therefore underestimates the regrowth. For example, if a tumour had decreased in size by 90%, then a regrowth of 10% is an increase of 100% from the perspective of the tumour at its smallest point. Ascertaining the specific tolerance of USS within this study remains difficult, being all of technology-, clinician-, and likely patient-dependent. Triplicate measurement may have allowed a statistical confidence to be approximated, though this remains non-trivial given the reality of the clinic, particularly for patients who are already undertaking non-essential tests as part of their therapy. Classification by USS therefore remains a delicate balance between assurance of correct class assignment and maximising study power, particularly for the less prevalent desensitised class. In essence, for this study it's been accepted that the likelihood of false-positives classification is increased in order to minimise false-negative classification.

The extended endocrine dataset's ability to inform future studies remains however, in (i) encouraging and directing the collection of sequential biopsies as an effective method of modelling acquired resistance and tumour dormancy, whether by less invasive and more feasible methods such as cell-free DNA (cfDNA) or circulating tumour cells (CTCs), and in (ii) revealing the potential role of epigenetic alteration in potentially inducing desensitisation to endocrine therapy.

Epigenetic modification resulting in desensitisation to endocrine therapy is an attractive hypothesis, providing a logical explanation as to the de/sensitivity 'switch' from responsive to resistant phenotypes as well as previous evidence supporting its potential role in a number of cancer types[246,247]. Moreover, DNA methyltransferase (DNMT) and histone deacetylase (HDACs) inhibitors have shown promise in breast cancer management, and in fact have demonstrated increased efficacy in combination[248,249], though benefit in solid tumours is not as well characterised as haematological malignancies[250]. In fact, in leukaemic patients, HDAC inhibition has been demonstrated to be time-dependent, supporting epigenetic modulation as a mechanism in acquiring resistance. Furthermore HDAC inhibition appears to be tumour selective, demonstrating synthetic lethality[251,252]. Perhaps most encouragingly, a phase-II trial (ENCORE 301) combined the HDAC inhibitor entinostat with exemestane to demonstrate both increased progression-free and overall survivals[253], resulting in entinostat being classified as an FDA Breakthrough Therapy for hormone receptor-

positive breast cancer and the development of an on-going phase-III trial (E21120)[254]. Epigenetic modification and alterations of chromatin structure may destabilise ER-responsive gene expression and force a switch to alternative ER-signalling pathways[255], and silencing of EREs via promoter methylation has been demonstrated *in vitro*[256]. A number of pioneer factors (*FOXA1*, *GATA3*, and *PBX1*) have been suggested to dynamically restructure chromatin and facilitate this alternative signalling[244,245,257]. *FOXA1* in particular has been associated with ER chromatin recruitment in metastatic disease[258] and is evident to show class-wise differential expression at extended treatment time points, warranting further analysis.

Ultimately HDAC inhibition stands as particularly attractive opportunity in explaining and managing acquired endocrine resistance, due to prior FDA approval for use in haematological malignancies and its logical potential as a resistance 'switch'. Direct validation of an epigenetic role in acquisition of resistance is clearly necessary, whether by IHC staining or more comprehensive sequencing approaches. Key to characterising this modulation is in tracking alterations over time as sequential biopsies and, in fact, epigenetic biomarkers in peripheral blood have shown promise in the early detection and characterisation of breast cancer[259,260].

5 | Continuous biomarker assessment by exhaustive survival analysis, *or* ***survivALL***

5.1 Context

Survival analysis comprises a collection of statistical procedures wherein the length of time leading up to an event is of interest. Within oncology, occurrence of and time to metastasis, local relapse or disease-specific death are frequently used to compare the relative survivals of two or more distinct populations, for example treated vs. placebo arms within a clinical trial. By demonstrating that a given event occurs later *and/or* less frequently within the treatment arm is evidence in favour of that same treatment. Whilst qualitative biomarkers can easily stratify cohorts into two or more groups for this purpose, the question of how to divide a cohort using a quantitative measure is considerably more challenging, where an obvious division is not apparent. Within a clinical context, dividing patients to allow a tractable decision is important, where reduction to binary choices of "treat" or "do not treat" are relevant, even though a greater number of subgroups may exist in reality. For instance, in the case of a prognostic signature outputting a continuous risk score, translation into whether a given patient should or should not be treated is key, though frequently an arbitrary division into two equal sized groups at a predefined level, such as the median, is employed.[i] However, this arbitrary-split approach is ignorant to the biomarker's distribution, as well as the composition of the cohort and any potential confounding clinical factors. In essence, median stratification of a patient cohort, based on the expression of an individual gene, would assume that for every high-risk patient there is an equivalent low-risk patient. The heterogeneous reality of patient cohorts makes assessing the robustness of a biomarker across several datasets problematic, with arbitrary dichotomisation unlikely to produce consistently reliable results due to random differences in sampling. In

[i] Though other arbitrary splits, such as comparing upper vs. lower quartiles, are common median dichotomisation will be highlighted throughout this chapter.

short, identifying a significant association with outcome by median separation in cohort A is unlikely to directly translate to a significant median separation in cohort B. The ultimate risk here is that an otherwise functional biomarker may be deemed irrelevant.

The opportunity to evaluate quantitative biomarkers *in silico* is in principle becoming easier, with increasing numbers of publicly available high-throughput molecular profiling datasets also providing patient follow-up data. In practice however, these datasets derived from cohorts of primary patient tissue are molecularly heterogeneous in their compositions[32], and are likely to exhibit varying proportions of multiple clinical factors associated with outcome (e.g. molecular subtypes, age, stage, grade and length of follow-up). It's therefore clear that this variation must be considered if a biomarker is to be successfully applied and validated in a meta-analysis, though these factors may be of unknown origin and difficult to correct in a multivariate analysis.

Building upon previous best-of-split methods[261–263], I developed *survivALL* (https://CRAN.R-project.org/package=survivALL) an R package to exhaustively calculate and visualise hazard ratios (HRs) for all possible points-of-separation and assess the association between a continuous measure and survival. As open-source software, *survivALL* allows for reproducible exhaustive survival analysis using public or independent datasets, in a highly transparent, automatable and extensible manner within the wider R package landscape. Complimenting *survivALL* is a companion web-based app — *survivAPP* (pearcedom.shinyapps.io/survivapp/) — allowing non-programmatic, drag-and-drop *survivALL* use. *survivALL* offers researchers the ability to perform exhaustive survival analysis highly relevant to the current state of –omics research.

In practice, *survivALL* computes the hazard ratio (HR)[264] for every possible point-of-separation,[i] allowing the magnitude — and frequency of — significant points-of-separation to be identified. Robust significance is determined using a non-parametric and dataset-specific bootstrap, performed as a 10,000-fold repeated calculation of HRs for random sample orderings, producing a distribution of expected/random HRs for each individual point-of-separation, to which observed true biological HRs are compared and significance calculated. Dataset-specific best points-of-separation are then optimised by maximising the desirability[265] of a combined *p*-value and absolute HR magnitude. To avoid Type-II errors associated with the multiple comparisons inherent in *survivALL*'s methodology, all examples are preceded by a single ancillary test of significance, performed using the biomarker as continuous predictor. Following this initial single test, all subsequent downstream analysis is performed using a combination of individual point-of-separation *p*-values and HR magnitude, and therefore no longer considers significance of an association with outcome, but rather where this association is most pronounced. This allows for robust determina-

[i] Hence, survival analysis + all possible points-of-separation = *survivALL*

tion of prognostic association whilst avoiding overcorrection in determining the optimum point-of-separation. *survivALL* is compared to arbitrary stratification approaches throughout the chapter and, whilst these are equivalent regardless of the chosen method (e.g. median, quartile separation), only median-dichotomisation is demonstrated for brevity.

5.2 Results

5.2.1 Exhaustive and arbitrary approaches to survival analysis

To illustrate the value of exhaustive survival analysis, I considered over-expression of human epidermal growth factor 2 (*HER2*/*ERBB2*), a well-established biomarker associated with poor prognosis in invasive breast cancer[38,266], in the largest single breast cancer gene expression dataset – METABRIC[43] (n = 1971 samples with complete disease-specific survival information). Importantly and unusually, METABRIC is split into two equally sized and composition-matched subsets, allowing for independent discovery and validation. Hazard ratio visualisation for all possible points-of-separation of ordered *ERBB2* expression within the METABRIC discovery cohort (n = 980) established 283/980 points-of-separation as significant (p < .05), confirming the expectation that increased *ERBB2* expression is associated with poor prognosis (Figure 5.2.1A). The span and location of these significant points are consistent with epidemiological evidence demonstrating that across the population ~20% of breast cancers overexpress HER2[267] and, importantly, none are located at the median. The contrasting Kaplan-Meier plots resulting from using the median (p = .48) and the data-driven most significant point-of-separation (p = 1x10^{-11}) for this dataset are also shown in Figure 5.2.1A.

The level of *ERBB2* expression associated with the discovery cohort-derived most significant point-of-separation was then applied to the validation cohort to predict its own specific most desirable point-of-separation (n = 991, Figures 5.2.1B). Prediction accuracy was measured as the number of patients alternatively classified compared to the validation cohort's own most desirable point-of-separation. This same measure of accuracy was also calculated using a median approach for comparison (Figure 5.2.1B), before application to the entire available transcriptome (19,628 genes, Figure 5.2.1C). As expected, *survivALL* significantly outperforms the median approach (p = 2x10^{-16}, Kolmogorov-Smirnov test), demonstrating the ability of exhaustive survival analysis to robustly determine prognostic classification in two suitably relatable datasets, without *a priori* knowledge of a gene's distribution within a population.

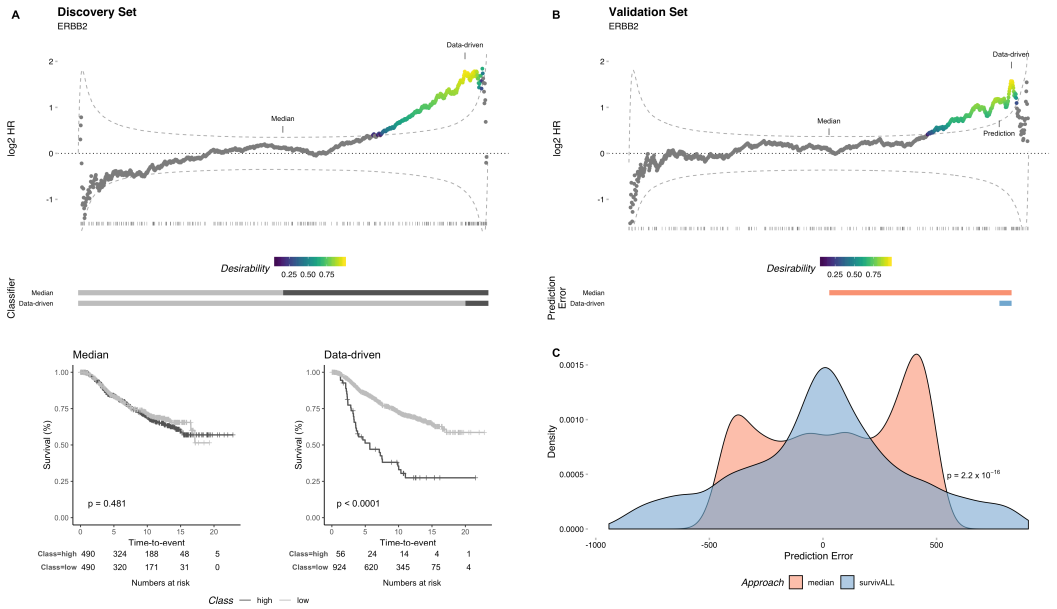

Figure 5.2.1: Exhaustive survival analysis of a continuous biomarker. (A) The plotALL() function allows hazard ratios (y-axis) and bootstrap significance ($p < 0.05$ colour scale; $p \geq 0.05$ grey) for all possible quantitative biomarker cut-points to be examined. Patients are ordered by increasing expression of the quantitative marker – in this example, *ERBB2* expression. Events (distant metastasis-free survival) are indicated as vertical lines below the plot, with darker colours indicating chronologically earlier events. Kaplan-Meier plots for *ERBB2* stratified using median ($n = 490$ vs. 490) and *survivALL* ($n = 924$ vs. 56) approaches are shown, comparing patients with expression lower (grey) or higher (black) than the points-of-separation determined in A. (B) Comparison of *survivALL* and median approaches in predicting optimal cohort separation for *ERBB2*. The expression values associated with both median and the dataset-specific most significant points-of-separation derived from the discovery set in A is applied to the validation dataset. Error (median = orange, *survivALL* = blue) is calculated as the number of patients who would be incorrectly assigned using these predictions, in relation to the validation cohorts own best separation. (C) Distribution of prediction error for all genes. *survivALL* significantly ($p = 2x10^{-16}$) outperforms a median approach in terms of predictive accuracy.

5.2.2 Reproducible determination of prognostic stratification using publicly available datasets

Real-world datasets rarely have similar compositions to the degree of the METABRIC discovery and validation cohorts, and it remained to determine if *survivALL* could perform reliable prognostic stratification in more dissimilar and noisier datasets. To test this, semi-random sub-samplings of the entire METABRIC dataset were simulated for three pre-defined proportions of estrogen receptor alpha (ER) positivity — 25, 50 & 75%. Using these dramatically variable compositions, the extent to which *survivALL* was able to track these differences in terms of the dataset-specific best point-of-separation for *GATA3*, a prominent luminal biomarker[244] was determined (Figure 5.2.2A). It was evident that as

the proportion of ER-positive and ER-negative samples shifted, the *survivALL* plots in turn shift in response, with the most significant point-of-separation consistently falling at the division between ER-positive and ER-negative samples. Importantly, although the most significant point-of-separation shifts between datasets, the corresponding level of expression that defined the ER+/- boundaries remained consistent (Figure 5.2.2B), indicating that *survivALL* is robustly determining a reliable and reproducible level of *GATA3* expression with which to stratify these patient cohorts, even across highly compositionally dissimilar datasets. In essence, this reproducible value represented a dataset-nonspecific or population best point-of-separation. Again, a median point-of-separation failed to define a level of expression that reproducibly stratified the patient cohorts and would therefore likely result in rejection of *GATA3* as a novel biomarker, contrary to the accepted literature (Figure 5.2.2B).

To further demonstrate the difficulty in deriving a reproducible point-of-separation across multiple datasets with variable compositions, I repeated this analysis using five publicly available, ER-positive tamoxifen treated datasets (GSE2990[65], GSE6532[207], GSE9195[206], GSE12093[208], GSE17705[209]; 5.1).

Dataset	*n*	ER+ %	PR+ %	HER2+ %	Node %	Event %	Event time (*range 0-16 years*)	Grade 1	2	3
GSE9195	77	100	77	0	47	13		14	20	24
GSE6532-p2	87	100	75	0	67	32		17	37	16
GSE6532-a	109	100	95	0	45	23		1	91	0
GSE12093	136	100	-	0	-	15		-	-	-
GSE2990	139	100	-	0	20	22		58	31	36
GSE17706	196	100	-	0	41	27		44	110	40

Table 5.1: ER-positive dataset compositions.

Each dataset was variable in its composition for several clinical factors including grade. It was therefore logical that each dataset, containing variable proportions of grade 1, 2, and 3 tumours would equally contain variable proportions of good and poor prognosis patients. Understandably any one arbitrary point-of-separation would fail to reflect this compositional variation, whereas exhaustive survival analysis could accurately reflect grade boundaries (Figure 5.2.3A). In assuming that tumour grade is an objective measurement that reflects prognosis equally for each of the 5 cohorts, each *survivALL*-derived dataset-specific best point-of-separation was therefore expected to relate to a consistent level of any biomarker correlated with grade. For this purpose, *AURKA*, a prominent biomarker of proliferation, being well explored and referenced in the literature for breast cancer, and known to be correlated with grade, was employed to explore the performance of a *known* biomarker[268,269]. As expected, *AURKA* was able to successfully determine GGI-determined

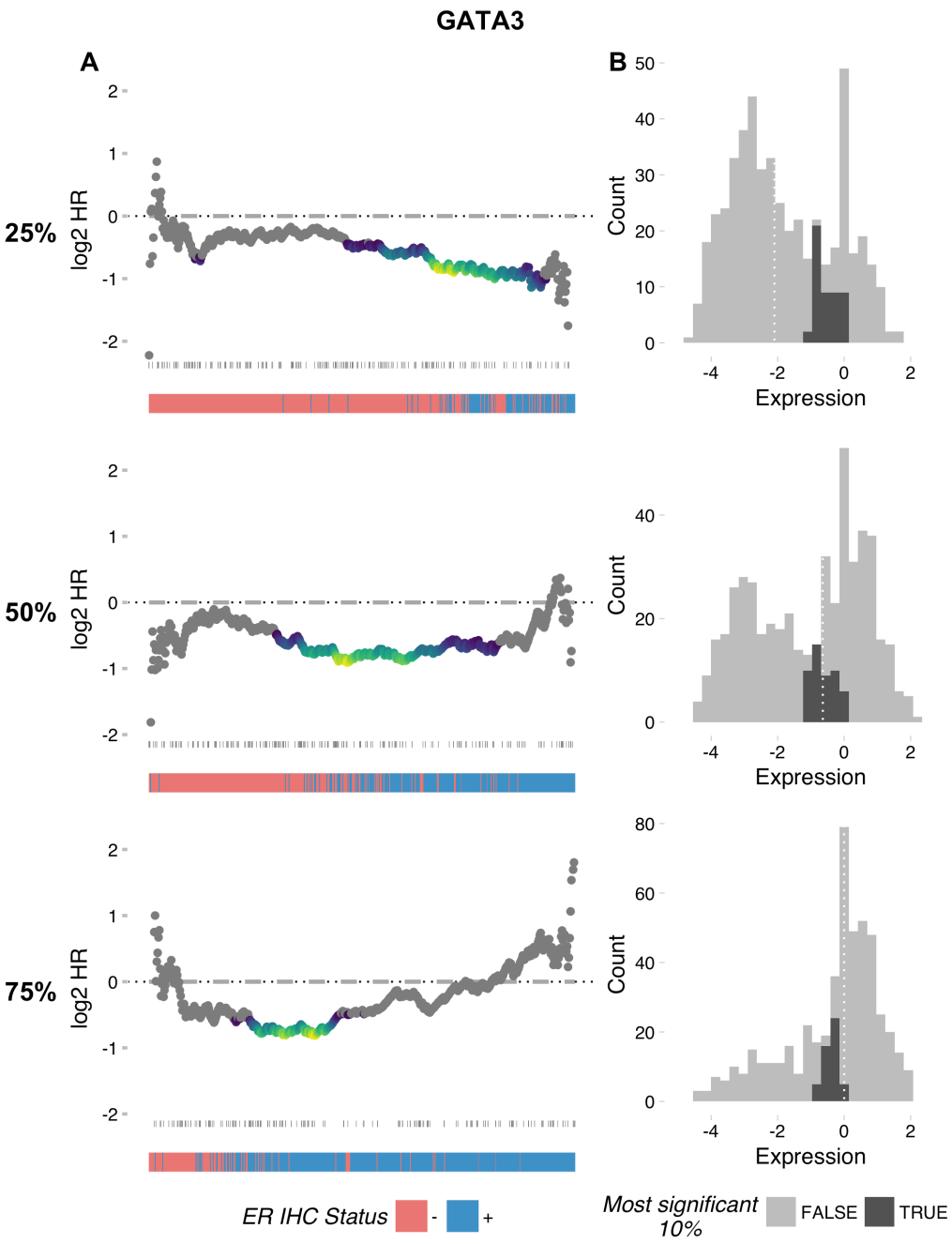

Figure 5.2.2: Determination of a reproducible cut-off for a continuous marker. (A) Analysis of METABRIC subset with variable composition. Three equally sized subsets (n = 500) with variable proportions of ER-positive samples (25, 50 and 75%) demonstrate the ability of *survivALL* to track the ER+/ER- (blue/red) boundary for *GATA3* expression, with the most significant point-of-separation (low-high significance = blue-yellow gradient) existing at this boundary for each subset. (B) Whilst the HR distribution changes with ER+ proportion, the 10% most significant points-of-separation consistently relate to the same level of *GATA3* expression. For comparison, the median level of expression (dashed white line) for each subset is also shown, which varies substantially between the three subsets.

grade boundaries across our 5 compositionally variable datasets where arbitrary methods would have otherwise failed (Figure 5.2.3B).

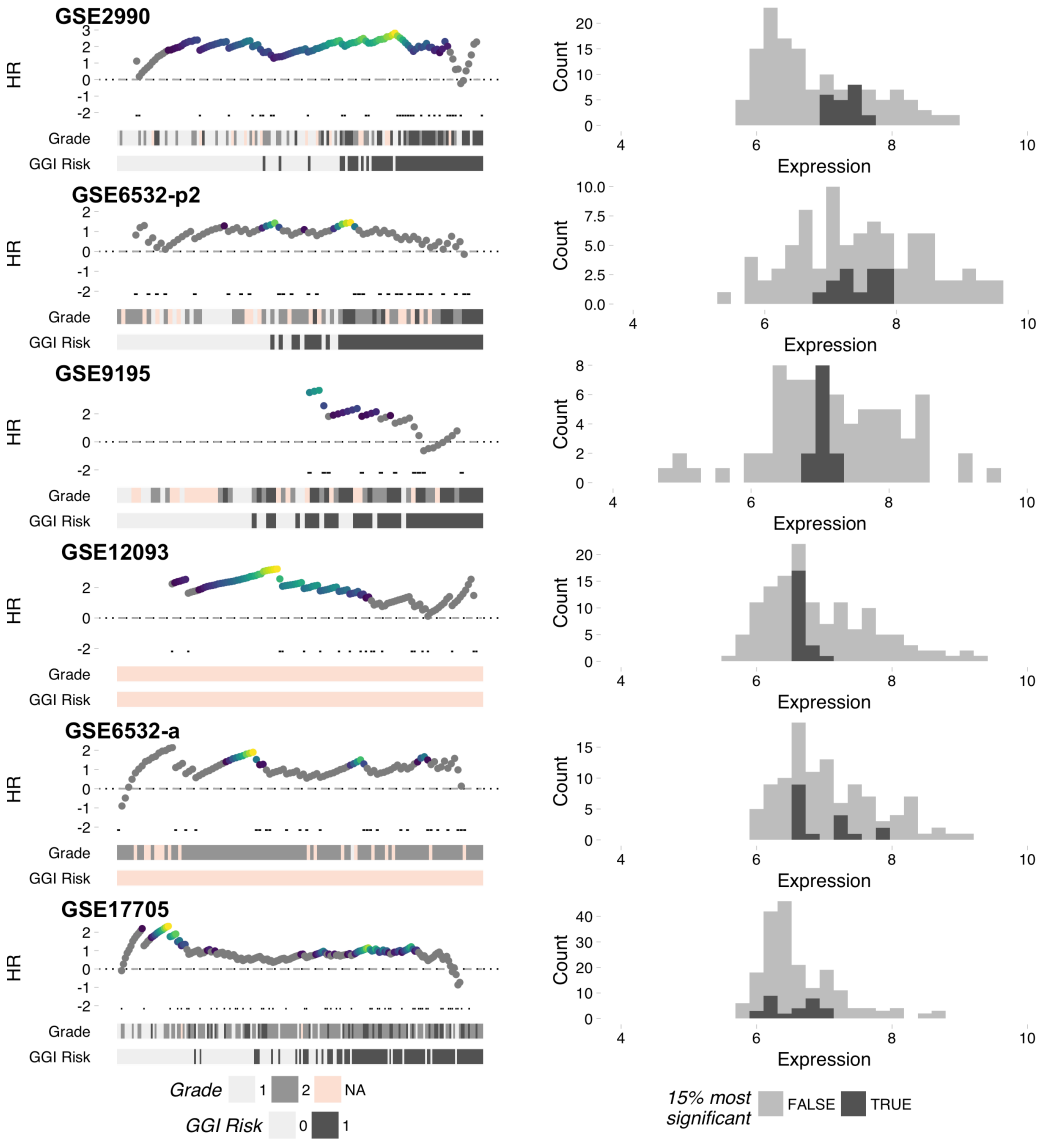

Figure 5.2.3: Continued analysis of compositionally variable datasets. Six ER-positive breast cancer, tamoxifen treated affymetrix datasets demonstrate *survivALL*'s ability to track the grade-defined boundaries (greyscale) for *AURKA* expression, with the most significant point-of-separation (less-more significant = blue-yellow) existing at these boundaries for each dataset. Whilst the HR distribution and dataset-specific best point-of-separation are evident to change with grade composition, the 15% most significant points-of-separation consistently relate to the same range of *AURKA* expression.

5.2.3 Significance does not guarantee a robust association with significance in meta-analysis

Whilst significant association of a biomarker can be determined in a meta-analysis using gene expression as a continuous predictor, this methodology does not reveal the direction of that association, i.e. good or poor prognosis. For a gene determined as significantly associated with prognosis in more than one dataset, the magnitude of these associations may be highly variable and even opposite (Figure 5.2.4). For the METABRIC discovery and validation cohorts this was evident to occur for ten genes, several of which (*ACY3*[270], *LRRK2*[271], *NUPR1*[272] & *UGT1A7*[273]) have been previously associated with cancer risk. This therefore represents a small but real quirk of survival analysis that must be considered in meta-analysis, emphasizing that caution should be employed when interpreting the results of quantitative biomarker evaluation to establish consensus findings.

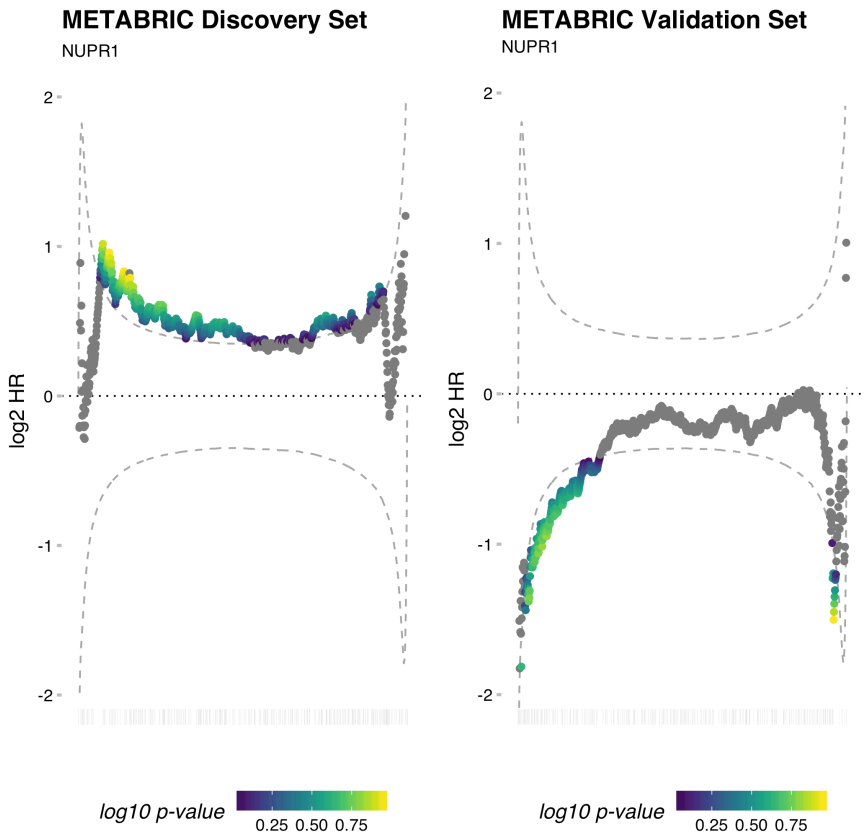

Figure 5.2.4: Biomarker significance and direction. *survivALL* reveals a gene, *NUPR1*, significantly associated with survival in both the discovery and validation METABRIC datasets, exhibiting opposite directions of association.

5.3 Perspectives

This chapter addresses the issues of assessing quantitative biomarkers for survival analysis, offering *survivALL* and *survivAPP* as tools to evaluate and overcome these challenges. Illustrated are situations relevant and common to researchers evaluating potential quantitative biomarkers in publicly available high-throughput datasets, using well established, but likely compositionally distorted examples. As an R package, *survivALL* allows greater resolution, transparency and flexibility compared to other online best-of-split tools, applicable to any public or proprietary dataset and usable in larger-scale automated data operations.

Researchers are increasingly using approaches such as KMplotter[274] and Cutoff Finder[275], either to median-split or highlight a dataset-specific best point-of-separation. This approach presents a number of potential problems highlighted in this chapter, though the single most notable problem with currently available survival analysis tools is the restriction of what data is available to be analysed – either alone or in combination as a meta-analysis – and the exact methods used. Furthermore, whilst a prognostic association can be considered using a continuous marker as the classifier itself, this ignores a number of informative factors, such as the direction and magnitude of association, as well as the optimal value to allow patient stratification into binary treatment groups. Many quantitative biomarkers and patient characteristics are associated with outcome. The distributions of these inter-related factors are likely to vary considerably between datasets, making the use of single arbitrary cut-points for stratification of prognostic groups highly context dependent. Our results demonstrate the value of evaluating all possible points-of-separation (using *survivALL*) to carefully consider whether and where to dichotomise a patient cohort.

Importantly, whilst the examples presented here relate to individual genes in breast cancer microarray datasets, *survivALL* is readily extensible to other diseases, data types or any other quantitative measure, including signatures or scores based upon marker combinations. Moreover, whilst this study focuses on dichotomisation, *survivALL* also allows for additional sub-populations with varying survivals to be visualised, as well as revealing potential confounding factors, producing multivariate analysis and demonstrating or uncovering an otherwise unknown factor when considering multiple datasets. Fundamentally, *survivALL* has been developed as open-source software, to flexibly integrate with other popular R packages, including the popular visualisation tool ggplot2 for customisable and scalable output. Most importantly, *survivALL* and survivAPP allows true biological effects and their relationship to survival to be revealed and reliably compared, within and between datasets, to move towards determining real-world clinically applicable biomarker cut-offs.

survivALL was published as a CRAN-distributed R package in 2017, detailing *survivALL* functions, examples and workflow in addition to the key rationale highlighted throughout this chapter. A pre-print article was made available to further detail the use and significance of *survivALL* and, similarly to this chapter, avoided reporting exploratory findings, instead using known biomarkers to illustrate the relative benefit of *survivALL* over arbitrary stratification approaches, where well-validated biomarkers are liable to be discarded. In this way, a *survivALL* user can be confident that, for a novel biomarker of unknown consequence, variation in composition between datasets, due to random sampling differences, will not affect the results, even with no *a priori* knowledge of how these confounding factors may manifest.

Consistent throughout the results here was a focus on dichotomisation, rationalised as the need for simple clinical decisions — i.e. treat or don't treat. It should be stressed, however, that the choice to analyse by dichotomisation removes information from a statistical analysis, potentially underpowering it. Key to *survivALL* is therefore its role as a complementary analysis, following careful initial assessment of a biomarker to determine a significant association with prognosis. Once statistical association has been satisfied, it is then appropriate to interrogate where patients may best be stratified by a given marker. Similarly, though dichotomisation has been highlighted, *survivALL* equally allows greater than two subpopulations to be interrogated. This is apparent in figure 5.2.4-left, where GSEs 6532-p2 and 17705 display a twin-peak of HRs that correlate with grade 1/2 and 2/3 boundaries. Trichotomisation (and beyond) using *survivALL* has not been robustly explored as part of this thesis, however. *survivALL*'s key utility is therefore in deriving these boundaries for patient stratification, to inform clinical application. That these dataset-specific points-of-separation can be tied back to consistent levels of expression (Figures 5.2.2 & 5.2.3) is highly encouraging. That a value even slightly consistent can be derived from datasets originating from multiple independent studies (purposefully left un-batch corrected) is even more encouraging still.

Though the nominal *survival* suggests a specific application in the life sciences, *survivALL* and exhaustive survival analysis are equally applicable beyond the remit of oncology and even biological science as a whole, extending to other data-types and, in theory, any continuous measure. However, whilst a logical assumption, this extended scope has as-of-yet not been demonstrated in practice.

6 | Conclusion

Within this thesis I have characterised temporal heterogeneity in breast cancer by two complimentary approaches — an examination of the role of tumour heterogeneity within the neoadjuvant window for untreated sequential biopsies; and the use of this same setting in characterising the molecular progression and regression of endocrine therapy and its resistance. Supplementing this was an additional methodological study detailing an improved method for appraising and comparing quantitative biomarkers. Taken together, these threads have illustrated the realities, needs, challenges and potential avenues for future research, which will be summarised here.

Neoadjuvant 'window of opportunity' studies make use of the pre-surgical nature of the neoadjuvant window to collect tumour material over time. Crucially, with the tumour still *in situ*, it is possible to molecularly profile these biopsies whilst accurately monitoring the clinical phenotype of the tumour and to synthesise these sources of information. It is important to remember that the ultimate aim for most of these studies is to characterise existing tumours to better inform our understanding of any given newly presenting cancer — both in terms of its likely response to therapy or its prognosis — as well as to pilot novel or existing therapies as part of a clinical trial, all within a greatly reduced time frame than would otherwise be possible. In the case of the latter, measurement of tumour volume, or markers of proliferation, are utilised to benchmark treatment efficacy and a reduction in size or proliferation (which to an extent can be considered synonymous) is evidence of response, correlating with long-term outcome. However, the use of protein or mRNA quantification of proliferative biomarkers, such as Ki67/*MKI67*, has long been confounded by potential variation resulting from normal tumour evolution. Cancers are known to be heterogeneous and they are known to vary in their composition over time but the use of *MKI67* as a measure of response to therapy ignores the potential that *MKI67*, along with any other gene, may vary independently of therapy. Similarly, and perhaps even more pertinent, is the idea that diagnostic classification may be changeable over the course of weeks or months, raising questions as to when clinical classification could and should be performed. It was therefore crucial to perform a control study characterising the variation between sequential tumour biopsies to both inform future, and corroborate past, studies.

Chapter 2 reports now published findings detailing 74 paired breast cancer biopsies from 37 patients taken at diagnosis and, between 14-53 days later, at surgery having received no-intervening treatment[1]. Perhaps unsurprisingly, though at odds with previous window study designs, variation was apparent between sample pairs taken from the same tumour. This variation was small relative to that observed in response to therapy, with a greatly reduced number of genes being significantly altered, but remained significant and consistent across patients nonetheless. Moreover, the apparent changes represented a number of genes previously reported as being involved in malignancy. This variation was equally apparent in 6 additional treated datasets, specifically for a selection of genes associated with response to surgical trauma and the induction of an early growth response. Importantly, diagnostic classification and proliferation were not observed to alter significantly between biopsies, suggesting that the use of *MKI67* as an outcome measure appropriate in the context of the neoadjuvant window.

The no-intervening treatment study results serve to inform the practicalities of sampling, highlighting the need to ensure that biopsies are performed in as uniform a manner as possible in order to minimise technical variation that would otherwise confound results. Moreover, they also strongly question the previously reported role of a core-needle biopsy in inducing a poor outcome associated immune response and supports their continued use as both a clinical and research tool[197]. That the information available within the neoadjuvant window can be maximised by considering changes over time has been demonstrated, and linking data from multiple biopsies to characterise tumour heterogeneity both highlights the importance of, and offers confidence for, the continued ethical use of sequential biopsies within the neoadjuvant window.

Perhaps lacking from Chapter 2 was conclusive proof that it was indeed the act of surgical excision that induced the characteristic early growth response, though this has fortunately been subsequently validated by another group[232], tying post-surgical excision biopsy X-ray duration with induction of many of the NIT signature genes. Similarly, it stands to reason that a score of treatment-unrelated affects could have been derived, to measure sampling-induced technical bias. This could find use in datasets where clinical data as to the exact protocol of how biopsies were taken and processed is missing. The long-term implications of the no-intervening treatment response will likely lie in future studies, wherein the NIT signature and dataset will serve as negative controls to which novel biomarkers can be compared and validated against.

In Chapter 3, a second cohort of patients were identified to characterise acquired resistance to endocrine therapy. Contrary to the NIT study, here patients with ER-positive disease underwent extended letrozole endocrine therapy. Patients were tracked over the course of years, and multiple sequential biopsies were collected over the course of treat-

ment. Our group has previously demonstrated the added benefit of an on-treatment treatment biopsy in tracking a dynamic response to therapy to classify ER-positive tumours as responsive or non-responsive[102]. However, despite initial responsiveness, a subset of tumours remain liable to develop resistance at a later date and, understandably, prediction of this acquired resistance phenotype was warranted. The extended duration of endocrine therapy throughout this study effectively allowed some tumours to stochastically become desensitised to treatment, whilst others remained sensitive and dormant, qualified by sequential measurements of tumour volume by USS. By transcriptomically comparing patients who did and who did not become desensitised to endocrine therapy the extended endocrine study attempted to first characterise whether patients could be separated based on their molecular characteristics before developing a classifier that could be applied for newly presenting patients. Indeed based on transcriptomic features, dormant and desensitised patients did display differential expression, manifesting as a reversion of an initially shared response to treatment, rather than two entirely independent genetic programs. This is perhaps the key factor that precluded an effective gene expression classifier being developed for acquired endocrine resistance, due to the relative similarity between classes, particularly at earlier on treatment time points where classification would be most desirable. However, despite this failure in developing a classification model, it remained apparent that significant differences in the expression of a number of epigenetic regulatory genes were repeatedly highlighted between classes. Though these genes remain to be conclusively validated, previous literature evidence, including an ongoing phase III trial[253,254] and the availability of FDA approved HDAC inhibitors encourage further scrutiny[276].

Though a generalisable classifier remains elusive, the extended endocrine study was able to highlight that, whatever changes are apparent, these are likely to emerge only after a length of time far beyond that of the normal neoadjuvant setting. The implications of this are that, if a predictive classifier could hypothetically be developed, there would likely be no tumour material left to test it on. Cell-free DNA and circulating tumour cells may offer a solution to this problem, wherein any residual tumour material is able to be assayed, in theory far beyond the time frame of the extended endocrine study. If desensitisation is influenced epigenetically, cfDNA may be the ideal candidate biomarker, relatively easily collected, stable and with previously work demonstrating the potential for epigenetic modifications from cfDNA to be identified in melanoma and colorectal cancers[277–279].

A major question that remains unresolved is whether desensitisation to endocrine therapy is through acquired alterations within cells under the conditions of treatment, or whether small subclonal compartments of the primary tumour exhibit resistance that is only revealed once a previously dominant clone is destroyed. Bulk tumour sampling, as was the practice for both neoadjuvant studies, sacrifices intra-tumour spatial heterogene-

ity for feasibility. However, purified as-of-yet unsequenced DNA exists for the extended endocrine dataset that could serve to characterise changes in subclonal composition. Emergence of a resistant clone is certainly an attractive hypothesis to explain desensitisation to treatment, though it could be argued that it poorly explains extended periods of tumour cell dormancy associated with late relapse in ER-positive disease. Single-cell RNA sequencing approaches have equally been demonstrated to illustrate the clonal dynamics of a tumour over time[280], though this more in-depth approach comes with associated increases in complexity and cost, as well as simply being a younger, less well validated method.

Though Chapter 4 characterises extended neoadjuvant endocrine therapy, it is not a perfect model. For instance, the time frame is notably less than the 5- or 10-years adjuvant endocrine therapy is commonly administered for, during which desensitisation to therapy can occur. The study duration, though extended, therefore does not characterise potential longer-term variation between dormant and desensitised patients. This is especially pertinent considering that endocrine response genes appear to deviate between classes more as time-on-treatment increases, even for extended time point samples. Moreover, acquisition of resistance in the real world would normally occur post-surgically, in the absence of the primary tumour, and the fact that the tumour remains *in situ* during the extended endocrine treatment study could be assumed to increase the likelihood of resistance occurring, given that there is simply a larger pool of cells and greater subclonal diversity for resistance to arise from.

Though validation of a predictive classifier and/or epigenetic involvement in desensitisation to therapy remains inconclusive, it should be noted that the Georgetown validation dataset was significantly smaller than the Edinburgh dataset and represents an underpowered comparison. However, criticism of the validation data should not be misunderstood as validation itself. It may prove that the Georgetown dataset could be better utilised in integration with the Edinburgh dataset, maximising sample size for a purely exploratory analysis and, indeed, the promise of similar extended neoadjuvant datasets may become available in the near future. Despite these issues, I believe that the results from Chapter 4, particularly those relating to the potential epigenetic influence on desensitisation to therapy, deserve further investigation. Having already generated the hypothesis using the entire transcriptome, validation could be achieved by smaller-scale approaches such as IHC. Furthermore, whilst able to suggest a role for epigenetic regulation, gene expression remains a poor way to quantify it, and alternative assays focussing on protein-chromatin associations, including chromatin immunoprecipitation (ChIP), may prove more sensitive, if combined with a long-term sequential monitoring program. Again the legacy of the extended endocrine therapy study lies both in its findings but also with the data itself, against

which findings from cell line and animal models can be compared and validated.

The neoadjuvant window is clearly a useful tool for the investigation of breast cancer. Whilst traditional neoadjuvant studies have focussed on assessing a response to treatment, this thesis highlights the interesting and complex phenotypes that can be effectively modelled, when combined with an expansive biobanking operation. Whilst breast cancer treatment is built upon well-established recommendations, the realities of treatment are varied and stochastic and, given enough patients and thoughtful planning, cohorts describing poorly characterised phenotypes can be assembled and clinical translatability maximised. Though, the benefits of tumour-derived over cell-line models are clear[281,282], and patient-derived tumour material represents the pinnacle of human translatability, in practice working with patient-derived material is tricky and the opportunity for technical biases is high. Chief amongst these is that for every non-ordinary cohort that is assembled, there may exist an unaccounted for confounding factor motivating its abnormality, for instance a common reason for delayed surgery combined with extended treatment, or indeed no treatment at all. Furthermore, the practical complexities are greatly increased. Chapter 2's no-intervening treatment highlighted the effect of biopsy method and sample processing can have on downstream data analysis, and there are many more opportunities for technical distortion to occur. This is particularly consequential when analysing sequential biopsies, where hybridisation failure of only one sample from a patient set can render the entire set unusable and sample size is disproportionately affected. Difficulties arising from the clinical, rather than laboratory, origin of neoadjuvant samples are countered by increasing recruitment and the use of paired statistical methods however. Whilst sequencing costs remain a bottle neck, investigations into human disease are likely to remain a compromise between reliability, relatability, and size and the reality remains that there is no quick or easy fix when it comes to cancer research — heterogeneity informs all stages of clinical and research practice, consistently questioning how best to investigate and compare different patients, or even tumour material from the same patient. Ultimately the results presented here encourage the ever growing acceptance of the role of heterogeneity in influencing and defining cancer research, highlighting the need for careful study design and analytical approach (as well as a dose of healthy scepticism), in order to translate and inform how heterogeneity should be managed in the clinic.

Acronyms and initialisms

Within the context of this thesis, reference to terms such as RNA, expression, and abundance are used interchangeably and in the context of mRNA expression and its quantification.

AI - aromatase inhibitor
ALTTO - adjuvant lapatinib and/or trastuzumab treatment optimisation
ATAC - arimidex, tamoxifen alone or combined
AURKA - aurora kinase A
CB - core-needle biopsy
cfDNA - cell-free DNA
ctDNA - circulating tumour DNA
CTC - circulating tumour cell
DNA - deoxyribonucleic acid
DNMT - DNA methyltransferase
EB - excision biopsy
EGFR - epidermal growth factore receptor
EMT - epithelial to mesenchymal transition
ER - oestrogen receptor
ERE - oestrogen response element
DGE(A) - differential gene expression (analysis)
FDA - US food & drug administration
FF - fresh frozen
FFPE - formalin-fixed and parafin-embedded
FIRST - fulvestrant first-line study comparing endocrine treatments
GSEA - geneset enrichment analysis
HDAC - histone deacetylase
HER2 - human epidermal growth factor receptor 2
HR - hormone receptor *or* hazard ratio
IHC - immunohistochemistry/chemical
IGF - insulin-like growth factor
IMPACT - immediate preoperative anastrozole, tamoxifen, or combined with tamoxifen
Lum - luminal
MAPK - mitogen-activated protein kinase
MAQC - microarray quality control
METABRIC - molecular taxonomy of breast cancer international consortium
NIT - no-intervening treatment

NCBI GEO - national centre for biotechnology information gene expression omnibus
NKI - national kaker instituut
NPI - nottingham prognostic index
NPV - negative predictive value
PCA - principle component analysis
PCR - polymerase chain reaction
PCNA - proliferating chain nuclear antigen
PDX - patient-derived xenograft
POETIC - peri-operative endocrine therapy for individualising care
PPV - positive predictive value
PR - progesterone receptor
ROR - risk of recurrence
RS - recurrence score
RNA - ribonucleic acid
RT-PCR - reverse transcription polymerase chain reaction
SAGE - serial analysis of gene expression
SAM - significance analysis of microarrays
SERD - selective oestrogen receptor downregulator
SERM - selective oestrogen receptor modulators
SSP - single sample predictor
TAM - tumour-associated macrophage
TNBC - triple-negative breast cancer
t-SNE - t-distributed stochastic neighbour embedding
UHRR - universal human reference RNA
USS - ultrasound sonography
WHO - world health organisation

Appendix 1: Tables

No-intervening treatment

		No. patients
IHC status		
	ER+/HER2-	22
	ER+/HER2+	8
	ER-/HER2-	7
Size (mm)		
	$\leq$20	12
	21-50	21
	51-60	2
	NA	2
Grade (Elston-Ellis)		
	1	3
	2	15
	3	18
	NA	1
Node		
	-	19
	+	18
Age		
	$\leq$40	5
	41-50	5
	51-60	5
	61-70	8
	71-80	9
	>80	5

Table 6.1: No-intervening treatment dataset clinico-pathological features.

Gene	Mean log2 FC	Gene	Mean log2 FC
HBA2	-1.23	HBB	-1.17
GOLGA6A	0.08	TMEM255B	0.09
LCA5L	0.11	C20orf141	0.11
ZNF565	0.11	LRCH1	0.14
APOLD1	0.14	ABL2	0.14
FAM86FP	0.15	EDNRB	0.15 gg
FLYWCH1	0.15	ABCA6	0.16
ITSN1	0.16	WDFY2	0.16
SPDYE3	0.18	GOLGA8K	0.19
PTGS2	0.20	KLF6	0.22
RASA3	0.22	SLC2A3P2	0.22
ATF3	0.22	GPR183	0.22
KRTAP19-6	0.23	NR4A3	0.24
NPIPA3	0.24	SIK1	0.25
NR4A2	0.26	ZSWIM4	0.30
NRP1	0.31	LAMB1	0.36
SRGN	0.36	C8orf4	0.39
SGK1	0.45	EGR3	0.52
GEM	0.53	SLC2A3	0.59
RASD1	0.61	MEG3	0.64
JUN	0.75	RGS1	0.76
NR4A1	0.77	ZFP36	0.83
CYR61	0.97	RGS2	1.02
FOS	1.13	EGR1	1.48
DUSP1	1.56	FOSB	1.60

Table 6.2: No-intervening treatment signature geneset.

Table 6.3: No-intervening treatment patient-level information A

Sample name	Source	Treatment	Biopsy	Organism	Gender	Age	Inter-biopsy-interval-(days)	Grade	Biopsy-type	IHC-substype
X1.CB	primary breast tumour	none	CB	Homo sapiens	Female	36	41	3	CB	er
X1.EB	primary breast tumour	none	EB	Homo sapiens	Female	36	41	3	EB	er
X2.CB	primary breast tumour	none	CB	Homo sapiens	Female	55	20	2	CB	er
X2.EB	primary breast tumour	none	EB	Homo sapiens	Female	55	20	2	EB	er
X3.CB	primary breast tumour	none	CB	Homo sapiens	Female	67	30	3	CB	her2
X3.EB	primary breast tumour	none	EB	Homo sapiens	Female	67	30	3	EB	her2
X4.CB	primary breast tumour	none	CB	Homo sapiens	Female	94	27	1	CB	er
X4.EB	primary breast tumour	none	EB	Homo sapiens	Female	94	27	1	EB	er
X5.CB	primary breast tumour	none	CB	Homo sapiens	Female	73	30	3	CB	tripleneg
X5.EB	primary breast tumour	none	EB	Homo sapiens	Female	73	30	3	EB	tripleneg
X6.CB	primary breast tumour	none	CB	Homo sapiens	Female	44	38	1	CB	er
X6.EB	primary breast tumour	none	EB	Homo sapiens	Female	44	38	1	EB	er
X7.CB	primary breast tumour	none	CB	Homo sapiens	Female	59	29	2	CB	er
X7.EB	primary breast tumour	none	EB	Homo sapiens	Female	59	29	2	EB	er
X8.CB	primary breast tumour	none	CB	Homo sapiens	Female	37	20	3	CB	her2
X8.EB	primary breast tumour	none	EB	Homo sapiens	Female	37	20	3	EB	her2
X9.CB	primary breast tumour	none	CB	Homo sapiens	Female	39	29	3	CB	tripleneg
X9.EB	primary breast tumour	none	EB	Homo sapiens	Female	39	29	3	EB	tripleneg
X10.CB	primary breast tumour	none	CB	Homo sapiens	Female	65	17	3	CB	her2
X10.EB	primary breast tumour	none	EB	Homo sapiens	Female	65	17	3	EB	her2
X11.CB	primary breast tumour	none	CB	Homo sapiens	Female	47	24	2	CB	er
X11.EB	primary breast tumour	none	EB	Homo sapiens	Female	47	24	2	EB	er
X12.CB	primary breast tumour	none	CB	Homo sapiens	Female	77	22	2	CB	er

X12.EB	primary breast tumour	none	EB	Homo sapiens	Female	77	22	2	EB	er
X13.CB	primary breast tumour	none	CB	Homo sapiens	Female	47	23	2	CB	her2
X13.EB	primary breast tumour	none	EB	Homo sapiens	Female	47	23	2	EB	her2
X14.CB	primary breast tumour	none	CB	Homo sapiens	Female	41	23	3	CB	her2
X14.EB	primary breast tumour	none	EB	Homo sapiens	Female	41	23	3	EB	her2
X15.CB	primary breast tumour	none	CB	Homo sapiens	Female	64	27	2	CB	er
X15.EB	primary breast tumour	none	EB	Homo sapiens	Female	64	27	2	EB	er
X16.CB	primary breast tumour	none	CB	Homo sapiens	Female	64	36	2	CB	er
X16.EB	primary breast tumour	none	EB	Homo sapiens	Female	64	36	2	EB	er
X17.CB	primary breast tumour	none	CB	Homo sapiens	Female	61	21	2	CB	her2
X17.EB	primary breast tumour	none	EB	Homo sapiens	Female	61	21	2	EB	her2
X18.CB	primary breast tumour	none	CB	Homo sapiens	Female	83	15	3	CB	er
X18.EB	primary breast tumour	none	EB	Homo sapiens	Female	83	15	3	EB	er
X19.CB	primary breast tumour	none	CB	Homo sapiens	Female	55	28	3	CB	er
X19.EB	primary breast tumour	none	EB	Homo sapiens	Female	55	28	3	EB	er
X20.CB	primary breast tumour	none	CB	Homo sapiens	Female	43	25	2	CB	her2
X20.EB	primary breast tumour	none	EB	Homo sapiens	Female	43	25	2	EB	her2
X21.CB	primary breast tumour	none	CB	Homo sapiens	Female	60	13	2	CB	er
X21.EB	primary breast tumour	none	EB	Homo sapiens	Female	60	13	2	EB	er
X22.CB	primary breast tumour	none	CB	Homo sapiens	Female	70	33	3	CB	her2
X22.EB	primary breast tumour	none	EB	Homo sapiens	Female	70	33	3	EB	her2
X23.CB	primary breast tumour	none	CB	Homo sapiens	Female	50	26	NA	CB	er
X23.EB	primary breast tumour	none	EB	Homo sapiens	Female	50	26	NA	EB	er
X24.CB	primary breast tumour	none	CB	Homo sapiens	Female	79	17	3	CB	er
X24.EB	primary breast tumour	none	EB	Homo sapiens	Female	79	17	3	EB	er
X25.CB	primary breast tumour	none	CB	Homo sapiens	Female	62	19	2	CB	tripleneg
X25.EB	primary breast tumour	none	EB	Homo sapiens	Female	62	19	2	EB	tripleneg

X26.CB	primary breast tumour	none	CB	Homo sapiens	Female	81	20	3	CB	tripleneg
X26.EB	primary breast tumour	none	EB	Homo sapiens	Female	81	20	3	EB	tripleneg
X27.CB	primary breast tumour	none	CB	Homo sapiens	Female	50	24	2	CB	er
X27.EB	primary breast tumour	none	EB	Homo sapiens	Female	50	24	2	EB	er
X28.CB	primary breast tumour	none	CB	Homo sapiens	Female	79	35	3	CB	tripleneg
X28.EB	primary breast tumour	none	EB	Homo sapiens	Female	79	35	3	EB	tripleneg
X29.CB	primary breast tumour	none	CB	Homo sapiens	Female	74	29	3	CB	er
X29.EB	primary breast tumour	none	EB	Homo sapiens	Female	74	29	3	EB	er
X30.CB	primary breast tumour	none	CB	Homo sapiens	Female	73	30	3	CB	tripleneg
X30.EB	primary breast tumour	none	EB	Homo sapiens	Female	73	30	3	EB	tripleneg
X31.CB	primary breast tumour	none	CB	Homo sapiens	Female	70	29	3	CB	er
X31.EB	primary breast tumour	none	EB	Homo sapiens	Female	70	29	3	EB	er
X32.CB	primary breast tumour	none	CB	Homo sapiens	Female	28	32	2	CB	er
X32.EB	primary breast tumour	none	EB	Homo sapiens	Female	28	32	2	EB	er
X33.CB	primary breast tumour	none	CB	Homo sapiens	Female	82	28	3	CB	tripleneg
X33.EB	primary breast tumour	none	EB	Homo sapiens	Female	82	28	3	EB	tripleneg
X34.CB	primary breast tumour	none	CB	Homo sapiens	Female	83	33	2	CB	er
X34.EB	primary breast tumour	none	EB	Homo sapiens	Female	83	33	2	EB	er
X35.CB	primary breast tumour	none	CB	Homo sapiens	Female	74	52	3	CB	er
X35.EB	primary breast tumour	none	EB	Homo sapiens	Female	74	52	3	EB	er
X36.CB	primary breast tumour	none	CB	Homo sapiens	Female	38	20	2	CB	er
X36.EB	primary breast tumour	none	EB	Homo sapiens	Female	38	20	2	EB	er
X37.CB	primary breast tumour	none	CB	Homo sapiens	Female	65	53	1	CB	er
X37.EB	primary breast tumour	none	EB	Homo sapiens	Female	65	53	1	EB	er

Table 6.4: No-intervening treatment patient-level information B

Sample name	ER-Allred	PR-Allred	HER2	Size-mm	Molecule	Label	Platform
X1.CB	6	NA	2+fish-ve	47	rna	biotin	GPL6947
X1.EB	6	NA	2+fish-ve	47	rna	biotin	GPL6947
X2.CB	8	NA	1+	29	rna	biotin	GPL6947
X2.EB	8	NA	1+	29	rna	biotin	GPL6947
X3.CB	8	NA	3+	23	rna	biotin	GPL6947
X3.EB	8	NA	3+	23	rna	biotin	GPL6947
X4.CB	0	0	NA	NA	rna	biotin	GPL6947
X4.EB	0	0	NA	NA	rna	biotin	GPL6947
X5.CB	0	0	2+fish-ve	33	rna	biotin	GPL10558
X5.EB	0	0	2+fish-ve	33	rna	biotin	GPL10558
X6.CB	7	NA	2+fish-ve	38	rna	biotin	GPL10558
X6.EB	7	NA	2+fish-ve	38	rna	biotin	GPL10558
X7.CB	8	NA	2+fish-ve	29	rna	biotin	GPL10558
X7.EB	8	NA	2+fish-ve	29	rna	biotin	GPL10558
X8.CB	7	5	3+	14	rna	biotin	GPL10558
X8.EB	7	5	3+	14	rna	biotin	GPL10558
X9.CB	3	0	2+fish-ve	31	rna	biotin	GPL10558
X9.EB	3	0	2+fish-ve	31	rna	biotin	GPL10558
X10.CB	8	NA	3+	15	rna	biotin	GPL10558
X10.EB	8	NA	3+	15	rna	biotin	GPL10558
X11.CB	8	NA	0	65	rna	biotin	GPL10558
X11.EB	8	NA	0	65	rna	biotin	GPL10558
X12.CB	8	NA	1+	22	rna	biotin	GPL10558

X12.EB	8	NA	1+	22	rna	biotin	GPL10558
X13.CB	8	NA	3+	24	rna	biotin	GPL10558
X13.EB	8	NA	3+	24	rna	biotin	GPL10558
X14.CB	7	NA	3+	17	rna	biotin	GPL10558
X14.EB	7	NA	3+	17	rna	biotin	GPL10558
X15.CB	8	NA	1+	34	rna	biotin	GPL10558
X15.EB	8	NA	1+	34	rna	biotin	GPL10558
X16.CB	8	NA	2+fish-ve	16	rna	biotin	GPL10558
X16.EB	8	NA	2+fish-ve	16	rna	biotin	GPL10558
X17.CB	8	NA	2+fish+ve	14	rna	biotin	GPL10558
X17.EB	8	NA	2+fish+ve	14	rna	biotin	GPL10558
X18.CB	7	NA	1+	21	rna	biotin	GPL10558
X18.EB	7	NA	1+	21	rna	biotin	GPL10558
X19.CB	4	0	1+	20	rna	biotin	GPL10558
X19.EB	4	0	1+	20	rna	biotin	GPL10558
X20.CB	8	NA	3+	26	rna	biotin	GPL10558
X20.EB	8	NA	3+	26	rna	biotin	GPL10558
X21.CB	8	NA	1+	53	rna	biotin	GPL10558
X21.EB	8	NA	1+	53	rna	biotin	GPL10558
X22.CB	3	0	3+	20	rna	biotin	GPL10558
X22.EB	3	0	3+	20	rna	biotin	GPL10558
X23.CB	7	NA	NA	NA	rna	biotin	GPL10558
X23.EB	7	NA	NA	NA	rna	biotin	GPL10558
X24.CB	8	NA	2+fish-ve	19	rna	biotin	GPL10558
X24.EB	8	NA	2+fish-ve	19	rna	biotin	GPL10558
X25.CB	0	NA	2+fish-ve	49	rna	biotin	GPL10558
X25.EB	0	NA	2+fish-ve	49	rna	biotin	GPL10558

X26.CB	0	0	2+fish-ve	31	rna	biotin	GPL10558
X26.EB	0	0	2+fish-ve	31	rna	biotin	GPL10558
X27.CB	7	NA	0	42	rna	biotin	GPL10558
X27.EB	7	NA	0	42	rna	biotin	GPL10558
X28.CB	0	0	1+	47	rna	biotin	GPL10558
X28.EB	0	0	1+	47	rna	biotin	GPL10558
X29.CB	4	NA	0	36	rna	biotin	GPL10558
X29.EB	4	NA	0	36	rna	biotin	GPL10558
X30.CB	2	0	2+fish-ve	25	rna	biotin	GPL10558
X30.EB	2	0	2+fish-ve	25	rna	biotin	GPL10558
X31.CB	8	NA	2+fish-ve	22	rna	biotin	GPL10558
X31.EB	8	NA	2+fish-ve	22	rna	biotin	GPL10558
X32.CB	8	NA	2+fish-ve	20	rna	biotin	GPL10558
X32.EB	8	NA	2+fish-ve	20	rna	biotin	GPL10558
X33.CB	2	NA	0	18	rna	biotin	GPL10558
X33.EB	2	NA	0	18	rna	biotin	GPL10558
X34.CB	8	NA	2+fish-ve	12	rna	biotin	GPL10558
X34.EB	8	NA	2+fish-ve	12	rna	biotin	GPL10558
X35.CB	4	0	0	37	rna	biotin	GPL10558
X35.EB	4	0	0	37	rna	biotin	GPL10558
X36.CB	8	NA	0	21	rna	biotin	GPL10558
X36.EB	8	NA	0	21	rna	biotin	GPL10558
X37.CB	8	NA	0	11	rna	biotin	GPL10558
X37.EB	8	NA	0	11	rna	biotin	GPL10558

Extended endocrine therapy

		Desensitised	Dormant	q
Biopsy Type				0.87
	CB	41	75	
	EB	20	41	
T				0.24
	T1	12	12	
	T2	25	53	
	T3	7	5	
	T4	14	30	
	T5	0	1	
M				0.09
	M0	55	92	
	M1	0	7	
	MX	3	2	
Node				0.09
	-	32	74	
	+	26	30	
Grade				0.09
	1	3	16	
	2	30	73	
	3	20	24	
Allred				0.09
	6	0	3	
	7	17	14	
	8	44	99	
HER2-positive				0.87
	-	38	95	
	+	8	17	
Recurrence				0.09
	Event	33	87	
	No-event	20	24	

Table 6.5: Extended endocrine treatment class relationships by clinico-pathological features. No significant differences in are apparent between classes (Fisher's Exact test).

Table 6.6: Extended endocrine therapy Edinburgh cohort tumour volume by USS information.

Patient ID	Biopsy taken	Days post-treatment	Volume	Rel. volume
32	TRUE	-1	31.68	100
32	FALSE	51	6.78	21.4
32	FALSE	63	15.82	49.94
32	FALSE	910	3.21	10.13
50	FALSE	0	12.92	100
50	FALSE	44	0.79	6.11
50	TRUE	86	5.18	40.09
50	FALSE	184	7.72	59.75
50	FALSE	231	1.32	10.22
109	FALSE	0	13.94	100
109	FALSE	52	8.65	62.05
109	FALSE	101	9.44	67.72
109	FALSE	192	2.72	19.51
109	FALSE	283	8.71	62.48
120	FALSE	0	36.87	100
120	FALSE	45	16.67	45.21
120	TRUE	87	5.65	15.32
120	FALSE	126	5.26	14.27
120	FALSE	169	4.84	13.13
139	FALSE	0	19.13	100
139	FALSE	49	3.18	16.62
139	FALSE	92	4.82	25.2
139	FALSE	162	0.98	5.12
147	FALSE	0	3.74	100
147	FALSE	84	0.98	26.2
147	FALSE	140	2.01	53.74
147	FALSE	196	3.35	89.57
171	TRUE	-12	17.91	100
171	FALSE	44	6.73	37.58
171	FALSE	86	3.04	16.97
171	FALSE	205	6.21	34.67
181	FALSE	0	8.01	100
181	FALSE	45	3.04	37.95
181	FALSE	87	2.15	26.84
182	FALSE	17	4.91	100
182	FALSE	45	1.36	27.7
182	FALSE	80	1.14	23.22
182	FALSE	171	1.13	23.01
182	FALSE	371	0.17	3.46
182	FALSE	743	0.18	3.67

188	FALSE	-3	8.17	100
188	FALSE	42	7.83	95.84
188	FALSE	84	2.83	34.64
188	FALSE	175	3.87	47.37
188	FALSE	217	3.38	41.37
188	FALSE	441	3.22	39.41
195	FALSE	0	1.34	100
195	FALSE	43	0.11	8.21
195	FALSE	78	0.21	15.67
195	FALSE	148	0.13	9.7
214	TRUE	14	4.11	100
214	FALSE	47	4.52	109.98
214	FALSE	89	0.69	16.79
215	FALSE	0	4.98	100
215	FALSE	45	1.88	37.75
215	FALSE	108	0.44	8.84
225	FALSE	0	212.17	100
225	TRUE	88	124.79	58.82
225	TRUE	179	50.97	24.02
225	FALSE	333	50.97	24.02
226	FALSE	0	4.34	100
226	FALSE	45	2.39	55.07
226	TRUE	143	0.09	2.07
226	FALSE	434	0.15	3.46
226	FALSE	556	0	0
249	FALSE	0	394.37	100
249	FALSE	59	118.79	30.12
249	TRUE	108	75.73	19.2
249	FALSE	295	56.09	14.22
249	FALSE	393	65.42	16.59
251	FALSE	0	10.93	100
251	FALSE	52	1.42	12.99
251	FALSE	101	1.79	16.38
251	FALSE	204	1.42	12.99
268	TRUE	-11	338.71	100
268	FALSE	45	243.6	71.92
268	FALSE	87	195.33	57.67
268	FALSE	157	87.07	25.71
269	TRUE	-8	4.84	100
269	FALSE	45	2.27	46.9
269	FALSE	108	0.75	15.5
276	FALSE	0	1.72	100
276	FALSE	46	1.63	94.77
276	FALSE	98	0.83	48.26
279	FALSE	0	3.32	100

279	FALSE	45	1.81	54.52
281	FALSE	0	3.69	100
281	FALSE	52	1.18	31.98
281	FALSE	101	0.65	17.62
289	FALSE	0	4.71	100
289	FALSE	45	3.59	76.22
289	FALSE	91	0.92	19.53
289	FALSE	189	1.51	32.06
289	FALSE	364	0	0
291	FALSE	0	15.39	100
291	FALSE	28	4.91	31.9
291	FALSE	98	3.14	20.4
298	FALSE	0	4.71	100
298	TRUE	22	0.82	17.41
299	FALSE	0	5.12	100
299	FALSE	64	0.47	9.18
299	FALSE	91	1.14	22.27
299	FALSE	133	0.48	9.38
300	TRUE	-6	2.54	100
300	FALSE	42	1.28	50.39
300	TRUE	84	1.03	40.55
300	FALSE	182	1.26	49.61
302	FALSE	15	13.98	100
302	FALSE	50	7.48	53.51
302	FALSE	106	3.14	22.46
309	FALSE	0	2.09	100
309	FALSE	64	0.38	18.18
309	TRUE	156	0.35	16.75
311	FALSE	0	2.4	100
311	FALSE	42	1.92	80
311	FALSE	84	0.68	28.33
323	TRUE	-15	5.56	100
323	FALSE	50	2.46	44.24
323	FALSE	91	1.07	19.24
325	FALSE	0	8.8	100
325	FALSE	57	6.01	68.3
325	FALSE	100	4.98	56.59
325	FALSE	211	2.54	28.86
328	TRUE	-25	13.11	100
328	FALSE	42	3.92	29.9
328	FALSE	91	2.93	22.35
329	FALSE	0	2.27	100
329	FALSE	43	1.74	76.65
329	FALSE	85	0.69	30.4
330	FALSE	0	2.83	100

330	FALSE	42	1.38	48.76
330	FALSE	77	1.18	41.7
333	TRUE	22	19.11	100
333	FALSE	38	2.91	15.23
333	FALSE	78	0.99	5.18
334	FALSE	0	5.89	100
334	FALSE	66	0.38	6.45
334	FALSE	105	0.42	7.13
340	FALSE	0	5.26	100
340	FALSE	49	3.38	64.26
340	TRUE	91	2.71	51.52
340	FALSE	196	3	57.03
341	TRUE	-11	4.05	100
341	FALSE	64	2.64	65.19
341	TRUE	106	0.75	18.52
341	FALSE	302	1.13	27.9
343	FALSE	0	16.53	100
343	FALSE	63	3.08	18.63
343	FALSE	273	0.28	1.69
346	FALSE	0	3.39	100
346	FALSE	49	3.04	89.68
346	FALSE	98	1.63	48.08
346	FALSE	196	3.52	103.83
347	TRUE	-7	14.02	100
347	FALSE	88	7.08	50.5
350	FALSE	0	13.82	100
350	FALSE	64	1.26	9.12
350	TRUE	134	2.79	20.19
350	FALSE	225	3.97	28.73
355	FALSE	0	0.68	100
355	FALSE	57	0.25	36.76
357	FALSE	0	3.63	100
357	FALSE	63	2.08	57.3
357	FALSE	105	0	0
364	FALSE	0	60.85	100
364	FALSE	48	26.01	42.74
364	TRUE	90	20.98	34.48
369	TRUE	0	6.37	100
369	FALSE	43	2.46	38.62
369	FALSE	98	2.42	37.99
374	FALSE	0	3.38	100
374	FALSE	42	0.71	21.01
374	FALSE	91	0.26	7.69
378	TRUE	0	10.86	100
378	FALSE	42	3.18	29.28

378	TRUE	106	0.95	8.75
378	FALSE	204	1.32	12.15
381	FALSE	0	15.86	100
381	FALSE	42	15.39	97.04
381	TRUE	106	6.36	40.1
381	FALSE	239	4.82	30.39
381	FALSE	358	2.98	18.79
382	FALSE	0	1.66	100
382	FALSE	78	1.34	80.72
382	FALSE	120	0.4	24.1
393	FALSE	0	5.83	100
393	FALSE	85	3.58	61.41
393	FALSE	148	2.01	34.48
394	TRUE	-6	6.6	100
394	FALSE	42	4.15	62.88
394	FALSE	105	3.04	46.06
395	FALSE	0	1.21	100
395	FALSE	211	0.13	10.74
399	FALSE	0	4.85	100
399	FALSE	36	2.88	59.38
399	FALSE	182	2.08	42.89
399	FALSE	330	1.3	26.8
400	FALSE	0	2.93	100
400	FALSE	85	0.97	33.11
400	FALSE	183	1.13	38.57
405	FALSE	0	1.66	100
405	FALSE	36	0.31	18.67
405	FALSE	113	0.31	18.67
405	TRUE	253	0.18	10.84
409	FALSE	0	3.05	100
409	FALSE	50	0.69	22.62
409	FALSE	176	0.24	7.87
413	FALSE	0	1.41	100
413	FALSE	50	0.88	62.41
413	FALSE	92	0.57	40.43
416	FALSE	0	3.05	100
416	FALSE	78	2.28	74.75
416	FALSE	120	1.33	43.61
283L	FALSE	0	5.73	100
283L	FALSE	52	2.86	49.91
283L	FALSE	101	1.03	17.98
373R	FALSE	-1	1.61	100
373R	FALSE	41	0.88	54.66
373R	FALSE	90	0.34	21.12

Table 6.7: Extended endocrine therapy Edinburgh cohort patient-level information.

Patient ID	Sample ID	Biopsy	Time Treated	T	N	M	Grade	Allred	HER2	event.os	time.os
109	109-1	1	-8	T2	N0	M0	3	8	NA	1	1083
120	120-1	1	-8	T4D	N1	M0	3	8	NA	0	4244
139	139-1	1	-11	T3	NO	M0	2	7	0	0	4697
171	171-1	1	-12	T3	N1	M0	2	7	NA	1	369
181	181-1	1	-8	T2	N0	M0	2	7	1	1	1793
182	182-1	1	-8	T2	N0	M0	2	8	0	1	1681
188	188-1	1	-11	T2	N1	M0	2	8	0	1	1323
195	195-1	1	-8	T4	N1	M0	2	7	0	NA	NA
214	214-1	1	-12	T4b	N0	M0	2	8	0	0	2854
215	215-1	1	-11	T2	N0	M0	3	6	0	0	3640
225	225-1	1	-10	T4b	N0	M0	3	7	0	1	811
226	226-1	1	-11	T2	N0	M1	3	8	0	1	745
249	249-1	1	-8	NA	NA	NA	NA	8	0	1	1377
251	251-1	1	-11	T4	N0	M0	2	8	0	1	486
268	268-1	1	-11	T3	N2	M0	3	8	1	NA	NA
269	269-1	1	-8	T2	N1	M0	3	8	0	1	2470
276	276-1	1	-7	T2	N1	M0	3	8	0	1	2641
279	279-1	1	-15	T2	N0	M0	1	8	0	1	2660
281	281-1	1	-8	T2	N0	M0	2	8	0	1	1032
283L	283L-1	1	-11	T1	N0	M0	1	8	0	NA	NA
289	289-1	1	-11	T4	N0	M0	2	7	0	1	3126
291	291-1	1	-6	T2	N1	MX	3	7	0	1	1448
298	298-1	1	-13	T2	N1	M0	2	8	0	0	2997
299	299-1	1	-11	T2	N0	M0	2	7	0	1	613
300	300-1	1	-6	T1	N0	M0	3	8	1	0	2857
302	302-1	1	-7	T1	N0	MX	2	7	0	0	3081
309	309-1	1	-11	NA	NA	NA	2	8	0	1	2796
311	311-1	1	-6	T1	N0	M0	3	8	1	1	2800
32	32-1	1	-1	NA	NA	NA	NA	8	NA	NA	NA
323	323-1	1	-15	T2	N0	M0	1	8	1	0	2873
325	325-1	1	-11	T2	N0	M0	2	8	0	0	2755
328	328-1	1	-25	T2	N0	M0	3	8	0	0	2832
329	329-1	1	-18	T4b	N0	M0	2	8	0	0	2680
330	330-1	1	-15	T2	N1	M0	2	8	0	0	2811
333	333-1	1	-15	T2	NX	M0	2	8	1	0	2828
334	334-1	1	-15	T4b	N1	M0	2	8	1	0	2822
340	340-1	1	-8	T4b	N0	M1	1	8	0	1	1467
341	341-1	1	-11	T2	N1	M0	1	8	0	1	2582
343	343-1	1	-16	T3	N1	M0	2	8	0	0	2649
346	346-1	1	-11	T2	N0	M0	2	8	1	1	2425
347	347-1	1	-7	T2	N0	M0	2	8	0	0	2615
350	350-1	1	-11	T4	N2	M0	3	8	0	1	280

355	355-1	1	-11	T1	N0	M0	2	8	0	0	2001
357	357-1	1	-11	T2	N0	M0	1	7	0	1	2706
364	364-1	1	-12	T2	N0	M0	2	8	0	0	2123
369	369-1	1	0	T2	N0	M0	2	8	1	0	2342
373R	373R-1	1	-9	T1	N0	M0	NA	8	0	NA	NA
374	374-1	1	-11	T1	N0	M0	2	8	0	0	2356
378	378-1	1	0	T2	N1	M0	2	8	0	0	2155
381	381-1	1	-15	T4b	N0	M0	2	8	0	0	534
382	382-1	1	-11	NA	NA	NA	2	8	0	0	1705
393	393-1	1	-12	T2	N0	M0	1	8	0	0	1572
394	394-1	1	-6	NA	NA	NA	2	8	0	0	1937
395	395-1	1	-13	T2	N0	M0	2	7	0	0	1885
400	400-1	1	-11	NA	N0	NA	2	8	0	1	544
405	405-1	1	-11	T1	N0	M0	2	8	0	1	1705
409	409-1	1	-13	T2	N1	M0	3	8	0	0	1717
413	413-1	1	-12	T1	N0	M0	2	8	0	1	878
416	416-1	1	-8	T4	N0	M0	2	8	0	0	1673
50	50-1	1	-12	T4D	N1	M0	2	8	NA	1	824
109	109-2	2	20	T2	N0	M0	3	8	NA	1	1083
120	120-2	2	17	T4B	N1	M0	3	8	NA	0	4244
147	147-2	2	17	T4	N0	M0	NA	7	NA	NA	NA
171	171-2	2	16	T3	N1	M0	2	7	NA	1	369
214	214-2	2	14	T4b	N0	M0	2	8	0	0	2854
215	215-2	2	15	T2	N0	M0	3	6	0	0	3640
225	225-2	2	18	T4b	N0	M0	3	7	0	1	811
226	226-2	2	14	T2	N0	M1	3	8	0	1	745
268	268-2	2	17	T3	N2	M0	3	8	1	NA	NA
269	269-2	2	17	T2	N1	M0	3	8	0	1	2470
291	291-2	2	14	T2	N1	MX	3	7	0	1	1448
298	298-2	2	22	T2	N1	M0	2	8	0	0	2997
299	299-2	2	21	T2	N0	M0	2	7	0	1	613
309	309-2	2	15	NA	NA	NA	2	8	0	1	2796
311	311-2	2	14	T1	N0	M0	3	8	1	1	2800
328	328-2	2	14	T2	N0	M0	3	8	0	0	2832
329	329-2	2	14	T4b	N0	M0	2	8	0	0	2680
330	330-2	2	14	T2	N1	M0	2	8	0	0	2811
333	333-2	2	22	T2	NX	M0	2	8	1	0	2828
334	334-2	2	22	T4b	N1	M0	2	8	1	0	2822
340	340-2	2	21	T4b	N0	M1	1	8	0	1	1467
343	343-2	2	21	T3	N1	M0	2	8	0	0	2649
346	346-2	2	22	T2	N0	M0	2	8	1	1	2425
347	347-2	2	39	T2	N0	M0	2	8	0	0	2615
350	350-2	2	29	T4	N2	M0	3	8	0	1	280
355	355-2	2	24	T1	N0	M0	2	8	0	0	2001
357	357-2	2	38	T2	N0	M0	1	7	0	1	2706

364	364-2	2	13	T2	N0	M0	2	8	0	0	2123
369	369-2	2	14	T2	N0	M0	2	8	1	0	2342
373R	373R-2	2	20	T1	N0	M0	NA	8	0	NA	NA
374	374-2	2	21	T1	N0	M0	2	8	0	0	2356
381	381-2	2	14	T4b	N0	M0	2	8	0	0	534
382	382-2	2	28	NA	NA	NA	2	8	0	0	1705
393	393-2	2	42	T2	N0	M0	1	8	0	0	1572
394	394-2	2	13	NA	NA	NA	2	8	0	0	1937
395	395-2	2	14	T2	N0	M0	2	7	0	0	1885
400	400-2	2	29	NA	N0	NA	2	8	0	1	544
409	409-2	2	14	T2	N1	M0	3	8	0	0	1717
413	413-2	2	14	T1	N0	M0	2	8	0	1	878
416	416-2	2	28	T4	N0	M0	2	8	0	0	1673
50	50-2	2	16	T4D	N1	M0	2	8	NA	1	824
120	120-2e	3	87	T4B	N1	M0	3	8	NA	0	4244
225	225-2e	3	88	T4b	N0	M0	3	7	0	1	811
249	249-2	3	108	NA	NA	NA	NA	8	0	1	1377
300	300-2	2	84	T1	N0	M0	3	8	1	0	2857
340	340-2e	3	91	T4b	N0	M1	1	8	0	1	1467
341	341-2	3	106	T2	N1	M0	1	8	0	1	2582
343	343-2e	3	105	T3	N1	M0	2	8	0	0	2649
364	364-2e	3	90	T2	N0	M0	2	8	0	0	2123
378	378-2	2	106	T2	N1	M0	2	8	0	0	2155
381	381-2e	3	106	T4b	N0	M0	2	8	0	0	534
50	50-2e	3	86	T4D	N1	M0	2	8	NA	1	824
109	109-4	4	303	T2	N0	M0	3	8	NA	1	1083
120	120-4	4	184	T4B	N1	M0	3	8	NA	0	4244
139	139-4	3	184	T3	NO	M0	2	7	0	0	4697
147	147-4	4	219	T4	N0	M0	NA	7	NA	NA	NA
171	171-4	3	236	T3	N1	M0	2	7	NA	1	369
181	181-4	3	121	T2	N0	M0	2	7	1	1	1793
182	182-4	3	884	T2	N0	M0	2	8	0	1	1681
188	188-4	4	559	T2	N1	M0	2	8	0	1	1323
195	195-4	4	184	T5	N1	M0	2	7	0	NA	NA
214	214-4	3	141	T4b	N0	M0	2	8	0	0	2854
215	215-4	3	163	T2	N0	M0	3	6	0	0	3640
225	225-4	4	392	T4b	N0	M0	3	7	0	1	811
226	226-4	4	143	T2	N0	M1	3	8	0	1	745
249	249-4	4	436	NA	NA	NA	NA	8	0	1	1377
251	251-4	4	261	T4	N0	M0	2	8	0	1	486
268	268-4	4	209	T3	N2	M0	3	8	1	NA	NA
269	269-4	3	149	T2	N1	M0	3	8	0	1	2470
276	276-4	3	136	T2	N1	M0	3	8	0	1	2641
279	279-4	3	135	T2	N0	M0	1	8	0	1	2660
281	281-4	3	121	T2	N0	M0	2	8	0	1	1032

283L	283L-4	3	177	T1	N0	M0	1	8	0	NA	NA
289	289-4	3	415	T4	N0	M0	2	7	0	1	3126
291	291-4	3	170	T2	N1	MX	3	7	0	1	1448
298	298-4	4	163	T2	N1	M0	2	8	0	0	2997
299	299-4	3	170	T2	N0	M0	2	7	0	1	613
300	300-4	3	219	T1	N0	M0	3	8	1	0	2857
302	302-4	3	177	T1	N0	MX	2	7	0	0	3081
309	309-4	3	156	NA	NA	NA	2	8	0	1	2796
311	311-4	3	121	T1	N0	M0	3	8	1	1	2800
32	32-4e	3	149	NA	NA	NA	NA	8	NA	NA	NA
32	32-4	4	1366	NA	NA	NA	NA	8	NA	NA	NA
323	323-4	3	142	T2	N0	M0	1	8	1	0	2873
325	325-4	3	359	T2	N0	M0	2	8	0	0	2755
328	328-4	3	128	T2	N0	M0	3	8	0	0	2832
329	329-4	3	149	T4b	N0	M0	2	8	0	0	2680
330	330-4	4	205	T2	N1	M0	2	8	0	0	2811
333	333-4	3	209	T2	NX	M0	2	8	1	0	2828
334	334-4	3	128	T4b	N1	M0	2	8	1	0	2822
340	340-4	4	275	T4b	N0	M1	1	8	0	1	1467
341	341-4	4	387	T2	N1	M0	1	8	0	1	2582
343	343-4	4	380	T3	N1	M0	2	8	0	0	2649
346	346-4	3	233	T2	N0	M0	2	8	1	1	2425
347	347-4e	3	158	T2	N0	M0	2	8	0	0	2615
347	347-4	4	262	T2	N0	M0	2	8	0	0	2615
350	350-4e	3	134	T4	N2	M0	3	8	0	1	280
350	350-4	4	247	T4	N2	M0	3	8	0	1	280
355	355-4	3	121	T1	N0	M0	2	8	0	0	2001
357	357-4	3	188	T2	N0	M0	1	7	0	1	2706
364	364-4	4	232	T2	N0	M0	2	8	0	0	2123
369	369-4	3	191	T2	N0	M0	2	8	1	0	2342
373R	373R-4	3	183	T1	N0	M0	NA	8	0	NA	NA
374	374-4	3	121	T1	N0	M0	2	8	0	0	2356
378	378-4	3	349	T2	N1	M0	2	8	0	0	2155
381	381-4	4	373	T4b	N0	M0	2	8	0	0	534
382	382-4	3	167	NA	NA	NA	2	8	0	0	1705
393	393-4	3	219	T2	N0	M0	1	8	0	0	1572
394	394-4	3	141	NA	NA	NA	2	8	0	0	1937
395	395-4	4	366	T2	N0	M0	2	7	0	0	1885
399	399-4	4	359	T4	N3	M0	2	8	0	0	1863
400	400-4	4	350	NA	N0	NA	2	8	0	1	544
405	405-4	4	282	T1	N0	M0	2	8	0	1	1705
409	409-4	4	209	T2	N1	M0	3	8	0	0	1717
413	413-4	4	163	T1	N0	M0	2	8	0	1	878
416	416-4	4	170	T4	N0	M0	2	8	0	0	1673
50	50-4	4	232	T4D	N1	M0	2	8	NA	1	824

Table 6.8: Extended endocrine therapy Georgetown cohort patient-level information A.

Patient ID	Sample ID	Timepoint	Days treated	Age	Surgery performed	T	N	M	ER	Grade
109	109-1	diagnosis	-8	78	WLE	T2	N0	M0	8	3
109	109-2	on-treatment	20	78	WLE	T2	N0	M0	8	3
109	109-4	long-term	303	78	WLE	T2	N0	M0	8	3
120	120-1	diagnosis	-8	64	WLE	T4d	N1	M0	8	3
120	120-2	on-treatment	17	64	WLE	T4d	N1	M0	8	3
120	120-3	on-treatment	87	64	WLE	T4d	N1	M0	8	3
120	120-4	long-term	184	64	WLE	T4d	N1	M0	8	3
124	124-1	diagnosis	-6	73	WLE	T2	N0	M0	8	3
124	124-2	on-treatment	24	73	WLE	T2	N0	M0	8	3
124	124-3	on-treatment	94	73	WLE	T2	N0	M0	8	3
124	124-4	long-term	1019	73	WLE	T2	N0	M0	8	3
135	135-1	diagnosis	-8	63	WLE	NA	NA	NA	8	2
135	135-4	NA	NA	63	WLE	NA	NA	NA	8	2
146	146-1	diagnosis	-8	73	no surgery	NA	NA	NA	7	3
146	146-3	NA	NA	73	no surgery	NA	NA	NA	7	3
146	146-4	long-term	737	73	no surgery	NA	NA	NA	7	3
159	159-1	diagnosis	-11	88	WLE	T2	N0	M0	8	2
159	159-4	long-term	198	88	WLE	T2	N0	M0	8	2
163	163-1	diagnosis	-11	84	WLE	T1	N1	M0	8	1
163	163-2	on-treatment	10	84	WLE	T1	N1	M0	8	1
163	163-4	NA	NA	84	WLE	T1	N1	M0	8	1
166	166-1	diagnosis	-18	73	WLE	T2	N0	M0	7	3
166	166-2	on-treatment	17	73	WLE	T2	N0	M0	7	3
166	166-4	long-term	293	73	WLE	T2	N0	M0	7	3
184	184-1	diagnosis	-11	92	WLE	T4b	N0	M0	8	2
184	184-4	long-term	268	92	WLE	T4b	N0	M0	8	2
187	187-1	diagnosis	-6	81	no surgery	NA	NA	NA	8	2
187	187-4	long-term	1139	81	no surgery	NA	NA	NA	8	2
188	188-1	diagnosis	-11	85	WLE	T2	N1	M0	8	2
188	188-4	long-term	559	85	WLE	T2	N1	M0	8	2
195	195-1	diagnosis	-8	67	NA	T4	N1	M0	7	2
195	195-4	long-term	184	67	NA	T4	N1	M0	7	2
208	208-1	diagnosis	-11	84	WLE	T3	N1	M0	8	2
208	208-2	on-treatment	45	84	WLE	T3	N1	M0	8	2
208	208-3	on-treatment	87	84	WLE	T3	N1	M0	8	2
208	208-4	NA	NA	84	WLE	T3	N1	M0	8	2
226	226-1	diagnosis	-11	78	no surgery	T2	N0	M1	8	3
226	226-2	on-treatment	14	78	no surgery	T2	N0	M1	8	3
226	226-4	NA	NA	78	no surgery	T2	N0	M1	8	3
248	248-1	diagnosis	-22	88	mast	T4	N0	M0	8	3
248	248-3	long-term	266	88	mast	T4	N0	M0	8	3

248	248-4	NA	NA	88	mast	T4	N0	M0	8	3
251	251-1	diagnosis	-11	76	WLE	T4	N0	M0	8	2
251	251-4	long-term	261	76	WLE	T4	N0	M0	8	2
257	257-1	diagnosis	-6	61	WLE	T2	N1	M0	8	1
257	257-4	NA	NA	61	WLE	T2	N1	M0	8	1
268	268-1	diagnosis	-11	72	mast	T3	N2	M0	8	3
268	268-2	on-treatment	17	72	mast	T3	N2	M0	8	3
268	268-4	long-term	209	72	mast	T3	N2	M0	8	3
279	279-1	diagnosis	-15	73	WLE	T2	N0	M0	8	1
279	279-4	NA	NA	73	WLE	T2	N0	M0	8	1
283	283-1	diagnosis	-11	80	WLE	T1	N0	M0	8	2
283	283-4	long-term	177	80	WLE	T1	N0	M0	8	2
298	298-1	diagnosis	-13	69	WLE	T1	N0	M0	8	2
298	298-2	on-treatment	22	69	WLE	T1	N0	M0	8	2
298	298-4	long-term	163	69	WLE	T1	N0	M0	8	2
304	304-1	diagnosis	-8	83	WLE	T2	N0	M0	8	2
304	304-4	long-term	247	83	WLE	T2	N0	M0	8	2
312	312-1	diagnosis	-11	57	WLE	T2	N0	M0	8	3
312	312-2	on-treatment	14	57	WLE	T2	N0	M0	8	3
312	312-4	NA	NA	57	WLE	T2	N0	M0	8	3
313	313-1	diagnosis	-6	80	no surgery	NA	NA	NA	8	2
313	313-4	NA	NA	80	no surgery	NA	NA	NA	8	2
319	319-1	diagnosis	-8	59	WLE	T2	N0	M0	8	2
319	319-4	NA	NA	59	WLE	T2	N0	M0	8	2
32	32-1	diagnosis	-1	61	WLE	T4	N0	M0	8	3
32	32-3	long-term	149	61	WLE	T4	N0	M0	8	3
32	32-4	long-term	1366	61	WLE	T4	N0	M0	8	3
330	330-1	diagnosis	-15	71	WLE	T2	N1	M0	8	2
330	330-2	on-treatment	14	71	WLE	T2	N1	M0	8	2
330	330-4	long-term	205	71	WLE	T2	N1	M0	8	2
341	341-1	diagnosis	-11	79	WLE	T2	N1	M0	8	1
341	341-3	on-treatment	106	79	WLE	T2	N1	M0	8	1
341	341-4	long-term	387	79	WLE	T2	N1	M0	8	1
343	343-1	diagnosis	-16	63	WLE	T3	N0	M0	8	2
343	343-2	on-treatment	21	63	WLE	T3	N0	M0	8	2
343	343-3	on-treatment	105	63	WLE	T3	N0	M0	8	2
343	343-4	long-term	380	63	WLE	T3	N0	M0	8	2
343	343-5	NA	NA	63	WLE	T3	N0	M0	8	2
343	343-6	NA	NA	63	WLE	T3	N0	M0	8	2
347	347-1	diagnosis	-7	75	WLE	T2	N0	M0	8	2
347	347-2	on-treatment	39	75	WLE	T2	N0	M0	8	2
347	347-3	long-term	158	75	WLE	T2	N0	M0	8	2
347	347-4	long-term	262	75	WLE	T2	N0	M0	8	2
350	350-1	diagnosis	-11	89	WLE	T4	N2	M0	8	3
350	350-2	on-treatment	29	89	WLE	T4	N2	M0	8	3

350	350-3	long-term	134	89	WLE	T4	N2	M0	8	3
350	350-4	long-term	247	89	WLE	T4	N2	M0	8	3
353	353-1	diagnosis	-11	84	no surgery	NA	NA	NA	NA	2
353	353-2	on-treatment	21	84	no surgery	NA	NA	NA	NA	2
353	353-4	NA	NA	84	no surgery	NA	NA	NA	NA	2
395	395-1	diagnosis	-13	75	WLE	T2	N0	M0	7	2
395	395-2	on-treatment	14	75	WLE	T2	N0	M0	7	2
395	395-4	long-term	366	75	WLE	T2	N0	M0	7	2
400	400-1	diagnosis	-11	73	WLE	NA	N0	NA	8	2
400	400-2	on-treatment	29	73	WLE	NA	N0	NA	8	2
400	400-4	long-term	350	73	WLE	NA	N0	NA	8	2
405	405-1	diagnosis	-11	65	WLE	T1	N0	M0	8	2
405	405-4	long-term	282	65	WLE	T1	N0	M0	8	2
409	409-1	diagnosis	-13	60	WLE	T2	N1	M0	8	3
409	409-2	on-treatment	14	60	WLE	T2	N1	M0	8	3
409	409-4	long-term	209	60	WLE	T2	N1	M0	8	3
413	413-1	diagnosis	-12	84	WLE	T1	N0	M0	8	2
413	413-2	on-treatment	14	84	WLE	T1	N0	M0	8	2
413	413-4	long-term	163	84	WLE	T1	N0	M0	8	2
416	416-1	diagnosis	-8	84	WLE	T4	N0	M0	8	2
416	416-2	on-treatment	28	84	WLE	T4	N0	M0	8	2
416	416-3	NA	NA	84	WLE	T4	N0	M0	8	2
416	416-4	long-term	170	84	WLE	T4	N0	M0	8	2
428	428-1	diagnosis	-10	85	WLE	T2	N1	M0	8	3
428	428-2	on-treatment	15	85	WLE	T2	N1	M0	8	3
428	428-3	long-term	199	85	WLE	T2	N1	M0	8	3
434	434-1	diagnosis	-11	75	WLE	T2	N1	M0	8	2
434	434-2	on-treatment	22	75	WLE	T2	N1	M0	8	2
434	434-3	on-treatment	92	75	WLE	T2	N1	M0	8	2
434	434-4	long-term	261	75	WLE	T2	N1	M0	8	2
50	50-1	diagnosis	-12	78	WLE	T4b	N1	M0	8	2
50	50-2	on-treatment	16	78	WLE	T4b	N1	M0	8	2
50	50-3	on-treatment	86	78	WLE	T4b	N1	M0	8	2
50	50-4	long-term	232	78	WLE	T4b	N1	M0	8	2
64	64-1	diagnosis	-8	89	Mast	T4	N2	M0	7	3
64	64-4	NA	NA	89	Mast	T4	N2	M0	7	3

Table 6.9: Extended endocrine therapy Georgetown cohort patient-level information B.

HER2	No. nodes+	Freshfrozen
NA	0	TRUE
NA	0	TRUE
NA	0	TRUE
NA	1	TRUE

NA	1	TRUE
NA	1	TRUE
NA	1	TRUE
NA	0	TRUE
NA	0	TRUE
NA	0	TRUE
NA	0	TRUE
NA	2	TRUE
NA	2	TRUE
NA	NA	FALSE
NA	NA	FALSE
NA	NA	FALSE
NA	2	TRUE
NA	2	TRUE
NA	1	FALSE
NA	1	FALSE
NA	1	FALSE
NA	0	FALSE
NA	0	FALSE
NA	0	FALSE
neg	1	TRUE
neg	1	TRUE
neg	NA	FALSE
neg	NA	FALSE
pos	10	TRUE
pos	10	TRUE
neg	3	TRUE
neg	3	TRUE
neg	2	FALSE
neg	2	FALSE
neg	2	FALSE
neg	2	FALSE
neg	1	TRUE
neg	1	TRUE
neg	1	TRUE
neg	0	FALSE
neg	0	FALSE
neg	0	FALSE
neg	0	TRUE
neg	0	TRUE
neg	0	FALSE
neg	0	FALSE
pos	8	TRUE
pos	8	TRUE
pos	8	TRUE

neg	0	FALSE
neg	0	FALSE
neg	1	FALSE
neg	1	FALSE
neg	1	TRUE
neg	1	TRUE
neg	1	TRUE
neg	0	FALSE
neg	0	FALSE
neg	2	FALSE
neg	2	FALSE
neg	2	FALSE
neg	NA	FALSE
neg	NA	FALSE
neg	3	FALSE
neg	3	FALSE
NA	0	TRUE
NA	0	TRUE
NA	0	TRUE
neg	4	TRUE
neg	4	TRUE
neg	4	TRUE
neg	6	TRUE
neg	6	TRUE
neg	6	TRUE
neg	0	FALSE
neg	0	FALSE
neg	0	FALSE
neg	0	FALSE
neg	0	FALSE
neg	0	FALSE
neg	0	TRUE
neg	0	TRUE
neg	0	TRUE
neg	0	TRUE
neg	9	TRUE
neg	9	TRUE
neg	9	TRUE
neg	9	TRUE
neg	NA	FALSE
neg	NA	FALSE
neg	NA	FALSE
neg	0	FALSE
neg	0	FALSE
neg	0	FALSE

neg	0	TRUE
neg	0	TRUE
neg	0	TRUE
neg	0	TRUE
neg	0	TRUE
neg	3	TRUE
neg	3	TRUE
neg	3	TRUE
neg	0	TRUE
neg	0	TRUE
neg	0	TRUE
neg	0	TRUE
neg	0	TRUE
neg	0	FALSE
neg	0	TRUE
neg	2	FALSE
neg	2	FALSE
neg	2	FALSE
neg	1	FALSE
neg	1	FALSE
neg	1	FALSE
neg	1	FALSE
NA	8	TRUE
NA	8	TRUE
NA	8	TRUE
NA	8	TRUE
NA	6	FALSE
NA	6	FALSE

Appendix 2: Published papers

SCIENTIFIC REPORTS

OPEN

Received: 15 February 2016
Accepted: 17 June 2016
Published: 07 July 2016

Tumour sampling method can significantly influence gene expression profiles derived from neoadjuvant window studies

Dominic A. Pearce[1], Laura M. Arthur[1], Arran K. Turnbull[1], Lorna Renshaw[1], Vicky S. Sabine[1,2], Jeremy S. Thomas[1], John M. S. Bartlett[1,2], J. Michael Dixon[1] & Andrew H. Sims[1]

Patient-matched transcriptomic studies using tumour samples before and after treatment allow inter-patient heterogeneity to be controlled, but tend not to include an untreated comparison. Here, Illumina BeadArray technology was used to measure dynamic changes in gene expression from thirty-seven paired diagnostic core and surgically excised breast cancer biopsies obtained from women receiving no treatment prior to surgery, to determine the impact of sampling method and tumour heterogeneity. Despite a lack of treatment and perhaps surprisingly, consistent changes in gene expression were identified during the diagnosis-surgery interval (48 up, 2 down; *Siggenes* FDR 0.05) in a manner independent of both subtype and sampling-interval length. Instead, tumour sampling method was seen to directly impact gene expression, with similar effects additionally identified in six published breast cancer datasets. In contrast with previous findings, our data does not support the concept of a significant wounding or immune response following biopsy in the absence of treatment and instead implicates a hypoxic response following the surgical biopsy. Whilst sampling-related gene expression changes are evident in treated samples, they are secondary to those associated with response to treatment. Nonetheless, sampling method remains a potential confounding factor for neoadjuvant study design.

Gene expression profiling of carcinomas has been widely used for molecular subtyping and prognostic prediction[1–9] producing a diverse library of gene classifiers. However, these signatures may be limited by the particular dataset used to produce the signature[10] and by the inherent cellular heterogeneity of tumours and the practical considerations of how samples are collected – in short the heterogeneity of the cohort and the heterogeneity of sampling.

A more informative and considered approach when performing molecular studies is to use matched biopsies from the same patient, allowing for both inter-patient variation to be controlled and changes occurring within a given tumour or organism to be more accurately modelled[11–13]. Matched sample pairs coupled with careful cohort selection should ensure that changes related to a given drug represent the largest source of variation, avoiding any unwanted contribution from confounding factors known or unknown and allow for greater statistical power with a smaller sample size[14,15].

Acquisition of multiple biopsies from an individual patient has been simplified by the more common use of neoadjuvant therapy for breast cancer, an increasingly popular treatment option for initially large, inoperable or locally advanced breast tumours, as well as operable cancers susceptible to specific treatments[16]. Pre-operative treatment with chemotherapy, or endocrine therapy in ER+ disease, not only increases rates of breast conserving surgery[17], but also allows a unique *in vivo* observation of tumour response to treatment[12,18]. This so-called 'window of opportunity' permits sequential biopsies of the same cancer to be taken at different time points during the course of the pre-operative treatment, allowing assessment of molecular changes in the tumour long before clinical evidence of response can be determined[11,13]. This has allowed the molecular effects of an administered drug to be studied and enabled biomarkers predictive of response or resistance to therapy to be identified at an increased rate. Recently our own group has demonstrated the added value of additional on-treatment measurements of gene expression to characterise and accurately predict the response to treatment[13,19].

Whilst the benefits of this 'window of opportunity' approach are certainly attractive for translational research, matched samples have commonly been collected under the assumption that variation observed between pairs will have occurred as a result of treatment – i.e. the results apparent are due to the drug alone. A control group is often not included in these studies and those that do are commonly limited to a handful of samples ($n = 8$[20], $n = 15$[21]) or are confounded by concurrent treatments. Conversely, studies that have compared multiple biopsies from the same patient in lieu of treatment are limited to only a fraction of the total molecular repertoire, most often focussing on hormone receptor status by IHC and not full transcriptomic profiling[22–25]. Whilst oestrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (Her2) exhibit high concordance between sample pairs in these studies, the growing understanding of breast cancer as an increasingly heterogeneous and polygenic disease necessitates a high-throughput approach.

Previous work[26] that has utilised larger-scale assays (a panel of 147 cancer related genes) investigated molecular variation under conditions of no-intervening treatment (NIT) in 21 paired core needle biopsy (CB) and excision biopsy (EB) samples. Here the diagnostic core biopsy was implicated in initiating an immune response, hypothesised to then be detected in a later surgically extracted excision biopsy. Potential stimulation of tumour-associated macrophages (TAMs) in response to CB was also reported, itself associated with poor prognosis in human breast cancer, raising concerns of taking multiple repeated biopsies from the same patient, underlining the importance of considering the full repertoire of genetic expression under conditions of no treatment.

Here we present the largest dataset to-date of untreated patient-matched breast cancer samples to determine whether, and to what extent, sample pairs exhibit molecular heterogeneity independent of treatment, and what the implications of any variation are in terms of the interpretation of patient-matched gene expression profiling studies. We explore possible causes of consistent differential expression and whether these reflect a wounding or immune response to the first biopsy, a hypoxia- or stress-induced response following blood supply interruption[27] or the normal growth and evolution of tumours over time[28].

Methods

Patient selection. Paired diagnostic core biopsies and surgical excision biopsies were identified from 37 patients with a primary histologically confirmed invasive breast cancer that did not receive any preoperative or neoadjuvant treatment at the Western General Hospital, Edinburgh, between 2003 and 2011. All patients gave informed consent to be included in the study, which was approved by the Lothian Regional Ethics Committee (2001/8/80 and 2001/8/81) and we confirm can that all experiments were performed in accordance with relevant guidelines and regulations. A summary and complete clinicopathological characteristics of the patients diagnostic core biopsies are given in Supplementary Tables S1 and S2.

Sample collection, RNA processing and microarray hybridisation. Core biopsies were taken at diagnosis in all patients using a 14-gauge automated needle device. Multiple cores were taken per tumour and combined as individual samples. Surgical excision biopsies of breast tumour were collected between 13 and 53 days later (mean interval = 27.5 days). Samples were snap frozen in liquid nitrogen and stored at −80 °C before homogenisation and RNA extraction using the RNeasy Mini Kit with RNAse Free DNAse treatment (Qiagen). RNA quantity and quality was assured using a Nanodrop 2000 c spectrophotometer (Thermo Scientific). RNA was reverse transcribed and amplified using the WT-Ovation FFPE System Version 2 (NuGEN), purified using the Qiaquick PCR purification Kit (NuGEN), biotinylated using the IL Encore Biotin Module (NuGEN), and purified using the minElute Reaction Cleanup Kit (Qiagen). At each step RNA/cDNA quantity and quality was assured by repeat assessment with the Nanodrop 2000 c prior to advancing to the next stage. Labelled cDNA was hybridized to Illumina Human HT-12 version 3 and version 4 whole-genome expression bead arrays according to the standard protocol for NuGEN labelled samples. Data was extracted using GenomeStudio software (Illumina).

Data analysis. All analysis was conducted in R (http://www.r-project.org) using software packages available via CRAN (http://cran.r-project.org/) and Bioconductor (http://www.bioconductor.org/). Data was pre-processed using the *lumi* package[29]. Log2 transformation, quantile normalisation and summarisation was performed for all Illumina probe profiles. Probe expression information was extracted and detected probes were standardised, firstly by passing a detection p-value threshold (≤ 0.05) and then by being called present in ≥ 3 samples. This was carried out three times, once each for three processing batches that comprised the dataset. A common feature list was determined by those probes common to all three dataset batches before re-mapping to Ensembl gene sequences using the *biomaRt* package[30,31]. Multiple probes-per-Ensembl ID were resolved by mean averaging. Batch correction was performed using ComBat[32] to integrate processing batches for further analysis (Supplementary Fig. S2). Sample groupings were compared using Pearson product-moment correlations. All significances were calculated by a two-tailed Wilcoxon rank sum test and corrected for multiple testing (FDR), unless otherwise stated. Intrinsic subtype assignment was performed using the *genefu* package[33]. Pairwise differential gene expression was calculated using significance analysis of microarrays (SAM), part of the *siggenes* package[34]. Hierarchical clustering was performed on pairwise fold change expression values using a complete linkage method. SAM analysis of treated data was performed using a hundred 18-sample permutations, to fairly match sample size between subsets, with the intersecting significant genes being taken as a mean average. The raw and processed data from this study can be accessed from NCBI GEO under the accession GSE76728.

Results

Consistent treatment-independent gene expression changes between diagnostic and surgical paired samples. Unsupervised clustering using the 500 most variable genes across the patient-matched

[1]Edinburgh Cancer Research Centre, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK. [2]Ontario Institute for Cancer Research, Toronto, Ontario, Canada. Correspondence and requests for materials should be addressed to A.H.S. (email: andrew.sims@ed.ac.uk)

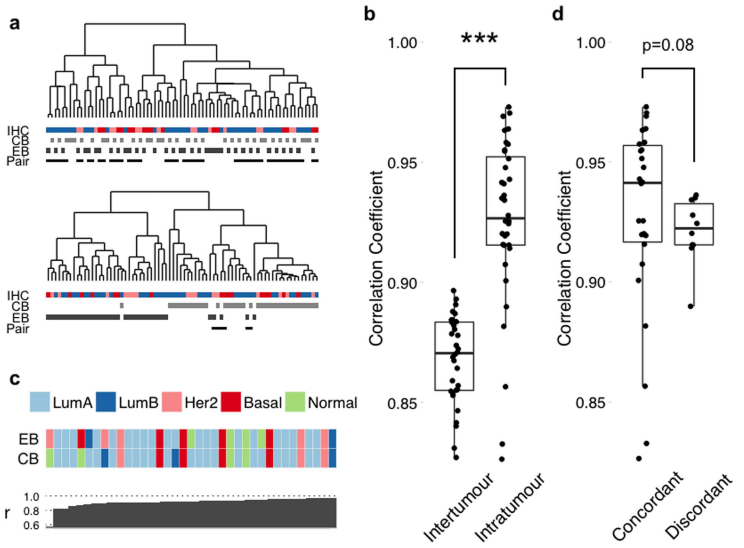

Figure 1. Evidence of treatment independent variation between breast cancer diagnostic core biopsies and surgical excision samples. (**a**) Hierarchical clustering of the 37 patient-matched diagnostic core and excision biopsy samples using the 500 most variable genes (upper) and a SAM derived signature of 50 genes consistently differentially expressed between core and excision biopsies (lower). Bars represent IHC status (ER+/Her2− = Blue; ER+/Her2+ = Pink; ER-/Her2− = Red) or biopsy method. Lower-most bar indicates where sample pairs have co-aggregated. Two-thirds (25/37) of the pairs cluster at the first level of the upper dendrogram, whereas pairwise association is lost in 31/37 cases for the lower. (**b**) There is a significantly stronger correlation between biopsy pairs (intra-tumour) than between different tumours (mean inter-tumour). ***$p < 0.001$. (**c**) Discordance in molecular subtype assignment between core and excision biopsies. Patients are ranked left to right by pairwise correlation. Colours represent SSP subtypes (Luminal A = Dark blue; Luminal B = Light Blue; Her2 = Pink; Basal = Red; Normal = Green). (**d**) Sample pairs were called discordant when Biopsy A ≠Biopsy B for at least 4/5 classifiers. Comparison of concordant vs. discordant pairwise correlations then revealed an inverse relationship between correlation and discordancy.

samples demonstrated high concordance, with 25/37 pairs observed to cluster at the first level of the dendrogram (Fig. 1a upper). In order to assess whether core biopsy and surgical excision sample pairs varied from one another, Pearson correlations were calculated between intra-tumour (paired) and inter-tumour (non-paired) samples (Fig. 1B). Intra-tumour differences were more significantly ($p = 7e\text{-}11$) correlated (median $r = 0.92$ range $r = 0.60–0.97$) than the mean inter-tumour variations (median = 0.87, range $r = 0.77–0.90$). Technical interference due to sample processing is expected to decrease correlation between samples, though only nominally[35]. Correlations below $r \simeq 0.97$ are likely due to underlying tumour heterogeneity, implicating a biological cause to the variation.

To determine the impact of differences in gene expression apparent between diagnostic core and excision biopsy pairs, pairwise SAM analysis identified 50 significantly differentially expressed genes (48 upregulated, 2 downregulated; FDR = 5%), between the surgical excision and diagnostic core biopsies (Supplementary Table S3). Clustering using these genes was able to separate samples by their sampling method in 31/37 cases (Fig. 1a lower). Five subtyping signatures (PAM50, Sorlie 2003, Hu 2006, Desmedt 2008, Wirapati 2008) were applied to each sample individually and concordant (subtype agreement between pairs with >3 subtype classifiers) vs. discordant (≤3 pair subtype agreements) results recorded (Fig. 1c (PAM50 only) and S5 (all methods)). Pairwise correlation between concordant and discordant assignments, revealed a trend ($p = 0.08$, Wilcoxon) between decreased pairwise correlation and greater discordance (Fig. 1d).

Pathway analysis of the 50 gene NIT signature revealed enrichment for MAPK signalling (DUSP1, JUN, NR4A1 and FOS), cancer specific (COX-2, PGE2, JUN and FOS), apoptosis induction (FOS and JUN) and genomic reformatting following (brain) ischaemia (EGR1 and JUN) pathways. A number of these genes are examples of 'early' or 'primary' growth response genes induced by both cell-extrinsic and cell intrinsic signals that do not require *de novo* protein synthesis for their expression[36].

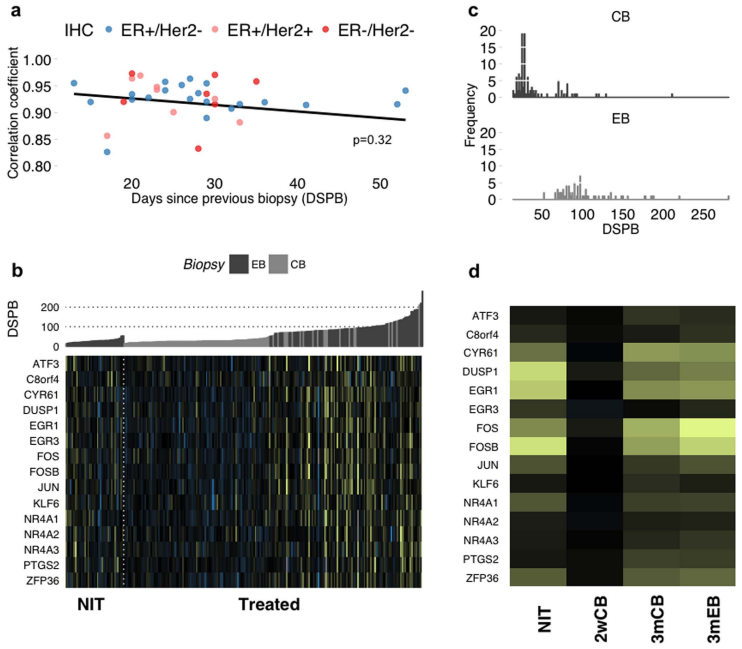

Figure 2. Factors associated with consistent gene expression changes between diagnostic core biopsy and surgical excision of breast tumours in the absence of treatment. (**a**) Pairwise correlations between biopsy pairs are not explained as either a function of time between biopsy ($p = 0.32$) or IHC status ($p = 0.43$). ER+/Her2− = Blue; ER+/Her2+ = Pink; ER-/Her2− = Red. (**b**) Heatmap showing differential expression of NIT signature genes in NIT and letrozole treated cohorts. Colours represent gene expression fold changes (up = yellow; down = blue) between samples and their subsequent patient-matched biopsies. Samples are ordered by increasing time between biopsies and reveal a pattern associated with either extraction method - CB (grey) or EB (dark grey) - or time. (**c**) Frequency distribution of biopsy time intervals. For further analysis, the letrozole treated data was split into three subsets – 2-week CB (2wCB), 3-month CB (3 mCB) and 3-month EB (3 mEB) – to investigate the effects of biopsy method and/or time on gene expression. (**d**) Mean expression fold changes since previous biopsy for 2 week, 3 month and NIT samples. 3 month subsets closely resemble NIT expression changes, though SAM analysis revealed a greater intersection of differentially expressed genes between NIT and 3 mEB samples than between NIT and 3 mCB samples.

Patient-matched gene expression changes may be associated with either time or biopsy methodology. Having determined significant and consistent changes in gene expression exist between diagnostic core and surgical excision pairs we sought to identify the underlying cause. Several hypotheses were immediately apparent – greater changes may occur following a shorter time interval between sampling if consistent gene expression changes reflect a wounding/immune response to the diagnostic core biopsy[26]; or it may be anticipated that expression patterns diverge over time to reflect tumour evolution; this may in turn be driven by tumour subtype[28]. However, comparisons of pairwise correlations defined by either IHC status subtype (ER+/Her2−, ER+/HER2+ or ER-/HER2−), PAM50 subtype (cross-table comparing IHC and PAM50 subtype assignment in Supplementary Table S4) or as a function of time interval between biopsies revealed no trend associated with either factor ($p = 0.43$ and $p = 0.32$ respectively) (Fig. 2a). This suggested that a progression in breast cancer-related biological changes were unlikely to be the root cause of the observed pairwise variation. To further investigate whether breast cancer biology could be responsible for the observed pairwise variation, we compared 7 breast cancer-related expression modules[37] between core and excision biopsies (Supplementary Fig. S3), which revealed no contrasting trend in gene expression. In conjunction to assigning cause, it remained equally important to determine whether our NIT signature genes were evident to alter in an equivalent treated data set. Using a patient-matched cohort treated with letrozole, collected and processed within our group in a manner analogous to the NIT data[13], there appeared to be a strong relationship between the time interval between biopsies and a

Neoadjuvant treatment	n	Biopsy Time Interval/days		Extraction Method			Dataset/Reference
		median	range	1st	On-	Final	
None	37	27	13–53	CB	–	EB	This study
Letrozole	122	107	13–884	CB	CB	CB EB	Turnbull *et al.*[13]
Celecoxib/none	22 15	*unspecified*	14–21	CB	–	EB	(Brandão *et al.*[21]
Anastrozole	81 On-18 Final	14	14–112	CB	CB	CB	(Smith *et al.*[57]
RAD001	21	14	14	CB	–	EB	(Sabine *et al.*[58]
Anthracycline-based Chemotherapy	69	*unspecified*	*unspecified*	CB	CB	EB	(Magbanua *et al.*[59]
None	56	*14*	*14*	CB		EB	Lopez-Knowles *et al.*[47]

Table 1. Composition of patient-matched neoadjuvant breast tumour datasets used within our study.

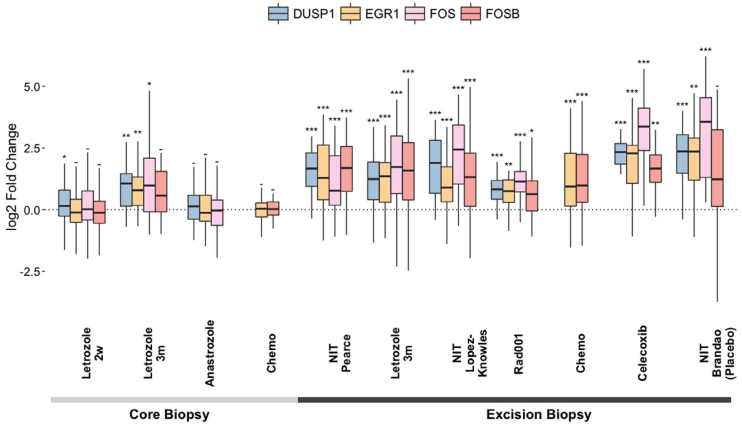

Figure 3. Multiple patient-matched datasets demonstrate shared changes in NIT early growth response genes. Pairwise analysis of four early growth response genes among the NIT signature in six validation datasets. These genes potentially represent an association between gene expression and sampling method, with surgically excised samples (EB) showing greater expression fold-change than their core biopsied (CB) counterparts. *$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$; $--$ = not significant.

subset of our NIT gene signature (Figs 2b and S4), with treated samples biopsied after 3 months exhibiting significant differential expression of these NIT genes. In these instances however, time interval almost wholly coincided with the final biopsy method and we therefore sought to determine which factor, if either, was dominant. Dividing the treated data into 3 subsets determined by the time since previous biopsy (Fig. 2c) as 2-week CB (2wCB, n = 95), 3-month CB (3 mCB, n = 18) and 3-month EB (3 mEB, n = 70), allowed comparison of changes assumed to result from treatment by both time interval and sampling method. Biopsy time interval was apparent as the factor most associated with the genes altered in our NIT data, with 3-month samples alone exhibiting expression fold changes similar to those observed in the NIT data, implying time on treatment as the defining factor (Fig. 2d). However, in tandem, global GSEA analysis was able to demonstrate that our NIT signature could significantly define the differences between untreated and treated data only when sampling method differed (NIT vs. 2wCB, p = 0; NIT vs. 3 mCB, p = 0.02; NIT vs. 3 mEB, p = 0.25). Similarly, only in the instance of excision biopsy was SAM analysis able to recapitulate >30% of the genes differentially expressed in the NIT data (mean = 47%). More so, out of 3955 differentially expressed treated-EB genes, seven of those common to NIT differential expression were amongst the top 15 in terms of fold change magnitude. Taken together, this implies both time on treatment as well as biopsy method are able to impact upon NIT signature gene expression.

Multiple patient-matched datasets also demonstrate changes in NIT early growth response genes. To further investigate the evident relationship between sampling method, time on treatment and NIT signature expression, we selected a panel of four genes (DUSP1, EGR1, FOS, FOSB) that were observed to be most representative of these factors, as well as being well characterised in the literature, for comparison in six external validation datasets (Table 1, Fig. 3). Significant differential gene expression was consistently observed to a greater degree only when an excision biopsy followed a previous core biopsy for these four genes.

Tumour sampling appears independent of an immune or wound-healing response. It remained important to determine whether gene expression changes in lieu of treatment were able to elicit either an immune or wound-healing response that may be detrimental to a patient's health or directly confound gene expression profiling results. Evaluation of both a published immune-related 9 gene panel, suggested to be upregulated in response to a diagnostic CB[26] and a 589 gene signature representative of a wound-healing response[38] failed to show any association with sampling method or time between biopsy, with sample expression unlikely to have been affected in terms of these biological categories (Supplementary Fig. S2).

Discussion

Our study reports gene expression changes in the largest cohort of sequential samples from patients receiving no-intervening treatment yet assembled to demonstrate the molecular variation that occurs independent of treatment in the neoadjuvant and preoperative setting. Significant pairwise changes in gene expression were observed and a 50 gene signature identified comprised of genes associated with a number of cell growth, cell stress and cancer related signalling pathways, including ATF3, EGR1, FOS, FOSB and JUN, each of which have been previously implicated in prognostic discrimination and pathogenesis of breast cancer[39–43] as well as other cancers[44].

We report that sampling method, specifically core versus excision biopsy, has a direct impact on gene expression and has the potential to introduce a confounding factor to downstream analysis. The most probable explanation of this expression variation between sample pairs is the technical issue of warm ischemia before newly biopsied samples are processed, when the cellular metabolic machinery attempts to mount a survival or apoptotic response before all metabolic activity ceases. Tissue ischemia may result from exclusion of the vascular supply or simply from handling by the surgeon, scrub nurse, pathologist and tumour bank personnel before the sample is frozen in liquid nitrogen. This time delay is likely (on average) to be significantly longer for surgical excision specimens than core biopsy samples. The effects of ischemia on gene expression has been described previously[45] and warm ischemia associated with the surgical extirpation of human tissues has significant effects on gene expression. These data support the careful monitoring of ischemic time for tissues harvested for the purpose of gene profiling. Similarly, Dash *et al.*[46] demonstrated significant changes in the expression of FOS, JUN, ATF3 in a study to examine the effect of processing time on prostate cancer samples.

Similar conclusions are proposed in a recent study published during the review process[47], where CB and EB pairs were also analysed in terms of correlation and differential gene expression. Variation between sample pairs was found to be evident though modest and those genes found to be differentially expressed intersected with those highlighted in this study (34% of our NIT signature genes). Of particular interest, the overlapping genes included the four genes highlighted in Fig. 3 as well as several others indicative of a stress or early growth response (RGS1, RGS2, ATF3, JUN), similarly proposed as a reaction to surgery-associated ischaemia. Indeed, the rationale of the study stemmed from a desire to investigate post-operative processing time and its effect on gene expression, citing two smaller studies which both highlighted ischaemia as a potential source of molecular variation[48,49]. Again, key molecules including FOSB were highlighted as being demonstrative of this effect. Importantly, the investigators additionally compared excision biopsies placed into RNAlater either immediately post-surgery or following an interval, and again observed early growth and stress response associated genetic expression to increase.

In a second analogous study, of 147 breast cancer-related genes measured by Nanostring in 21 patients, Jeselsohn *et al.*[26] proposed that sequential breast cancer biopsies reveal activation of an immune response, characterised by a panel of 9 immune-related genes, of which CD68 is known to activate tumour-associated macrophages and implicated in increasingly severe prognosis[50], as well as being present in the clinically available Onco*type* DX® breast cancer assay[8]. However, none of these genes were found to be significantly differentially expressed in our larger, whole genome cohort (Supplementary Fig. S1). Conversely, a number of studies have demonstrated a high degree of concordance between classical IHC markers for breast cancer, namely oestrogen, progesterone and HER2 receptors[24,25] between diagnostic core and excision biopsies, suggesting discordance may be limited to the level of transcription. Our study demonstrates that pairwise variation at the transcriptome level is not limited to the classical markers of breast cancer, though in some cases paired samples may be classified differently using molecular signatures (Figs 1c and S5). The causative factor behind this variation is most likely due to sampling method in our NIT cohort, with surgical resection resulting in gene expression in response to stress. It remains unclear whether this effect translates to samples following treatment, with time on treatment being observed to mediate molecular changes against the background of the sampling method.

It is important that results from preoperative window or neoadjuvant studies are carefully scrutinised, as a study by Morrogh *et al.*[51] - using a 502 cancer-related gene panel to examine 16 paired patient samples, 8 of which were untreated controls - claimed that window trials are influenced by the wound-healing process. They proposed that upregulation of MLL and FOSB was evidence of this, irrespective of treatment. This mirrored our own findings, though it is necessary to note that Morrogh *et al.* were limited by sample size and inconsistencies in the effect direction of proliferation markers (increased Ki67, but decreased PCNA). Nonetheless it remains likely that the overall pathway level message within our study - upregulation of proliferation - was a consequence of an early growth response to the biopsy itself. It is crucial to determine whether there is a genuine immune response associated with biopsy type as inflammatory signatures have been associated with poor anti-proliferative responses to aromatase inhibitors[13,52]. Clear evidence of both an immune or wound response signature[53] was completely absent in our untreated samples, implying that any contribution of biopsy methodology to an inflammatory response is likely to be minor.

With the potential for pairwise variation irrespective of treatment, our study raises potential concerns of the suitability of the neoadjuvant window in gene expression profiling studies. Recent results of the ALTTO (Adjuvant Lapatinib And/Or Trastuzumab Treatment Optimisation) clinical trial were, however, found to be consistent with the predicted benefits from the neoALTTO trial. This supported the utility of the neoadjuvant setting

as a suitable and important window for evaluating promising new targeted agents[54], as well as the continued use of patient-matched samples to assess intervention studies for translational research. Nonetheless it remains critical to understand that whilst patient-matched samples reduce variation due to individuals, all possible sources of variation must be considered for an optimal experimental design. For example, in an intervention study to assess dietary changes on normal breast tissue from pre-menopausal women it was considered optimal to schedule the sequential biopsies one menstrual cycle apart, rather than using a fixed window of time[55], as there is clear evidence that menstrual changes in oestrogen levels caused significant changes in gene expression[56]. Underlying tumour heterogeneity is an inevitable variable when performing neoadjuvant or window studies, however our study suggests that the method of sample collection should be considered along with treatment, time interval and clinicopathological features as an important potential confounding factor. These considerations are of particular importance if a study's purpose is the development of a prognostic/predictive classifier or identification of a biomarker, with genes present in our NIT signature excluded from the analysis.

Our study demonstrates that consistent molecular changes arise in tumours in the absence of treatment and these can impact upon classification. These changes appear to be an artefactual ischemic response resulting from the sampling methodology itself, rather than reflecting the effects of a previous biopsy. Careful consideration should be given in future studies that seek to illustrate molecular changes between paired biopsies in the neoadjuvant setting for breast cancer and likely other cancers that make use of the same experimental design.

References

1. Dekker, T. J. A. *et al.* Reliability of core needle biopsy for determining ER and HER2 status in breast cancer. *Ann. Oncol.* **24,** 931–937 (2013).
2. van de Vijver, M. J. *et al.* A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* **347,** 1999–2009 (2002).
3. Filipits, M. *et al.* A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. *Clin. Cancer Res.* **17,** 6012–6020 (2011).
4. Ma, X.-J. *et al.* A five-gene molecular grade index and HOXB13:IL17BR are complementary prognostic factors in early stage breast cancer. *Clin. Cancer Res.* **14,** 2601–2608 (2008).
5. Jankowitz, R. C. *et al.* Prognostic utility of the breast cancer index and comparison to Adjuvant! Online in a clinical case series of early breast cancer. *Breast Cancer Res.* **13,** R98 (2011).
6. Sotiriou, C. *et al.* Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J. Natl. Cancer Inst.* **98,** 262–272 (2006).
7. Toussaint, J. *et al.* Improvement of the clinical applicability of the Genomic Grade Index through a qRT-PCR test performed on frozen and formalin-fixed paraffin-embedded tissues. *BMC Genomics* **10,** 424 (2009).
8. Paik, S. *et al.* A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N. Engl. J. Med.* **351,** 2817–2826 (2004).
9. Sorlie, T. *et al.* Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc. Natl. Acad. Sci.USA* **100,** 8418–8423 (2003).
10. Ein-Dor, L., Kela, I., Getz, G., Givol, D. & Domany, E. Outcome signature genes in breast cancer: is there a unique set? *Bioinformatics* **21,** 171–178 (2005).
11. Sabine, V. S. *et al.* Gene expression profiling of response to mTOR inhibitor everolimus in pre-operatively treated post-menopausal women with oestrogen receptor-positive breast cancer. *Breast Cancer Res Treat* **122,** 419–428 (2010).
12. Sims, A. H. & Bartlett, J. M. Approaches towards expression profiling the response to treatment. *Breast Cancer Res.* **10,** 115 (2008).
13. Turnbull, A. K. *et al.* Accurate Prediction and Validation of Response to Endocrine Therapy in Breast Cancer. *J. Clin. Oncol.* **33,** 2270–2278 (2015).
14. Sims, A. H. Bioinformatics and breast cancer: what can high-throughput genomic approaches actually tell us? *J Clin Pathol* **62,** 879–885 (2009).
15. Turnbull, A. K. *et al.* Direct integration of intensity-level data from Affymetrix and Illumina microarrays improves statistical power for robust reanalysis. *BMC Med. Genomics* **5,** 35 (2012).
16. Untch, M., Konecny, G. E., Paepke, S. & von Minckwitz, G. Current and future role of neoadjuvant therapy for breast cancer. *Breast* **23,** 526–537 (2014).
17. Mamounas, E. P. Facilitating breast-conserving surgery and preventing recurrence: aromatase inhibitors in the neoadjuvant and adjuvant settings. *Ann. Surg. Oncol.* **15,** 691–703 (2008).
18. Macaskill, E. J. & Dixon, J. M. Neoadjuvant use of endocrine therapy in breast cancer. *Breast J.* **13,** 243–250 (2007).
19. Arthur, L. M. *et al.* Molecular Changes in Lobular Breast Cancers in Response to Endocrine Therapy. *Cancer Res.* **74,** 5371–5376 (2014).
20. Morrogh, M. *et al.* Differentially expressed genes in window trials are influenced by the wound-healing process: lessons learned from a pilot study with anastrozole. *J Surg Res* **176,** 121–132 (2012).
21. Brandão, R. D. *et al.* A randomised controlled phase II trial of pre-operative celecoxib treatment reveals anti-tumour transcriptional response in primary breast cancer. *Breast Cancer Res.* **15,** R29 (2013).
22. Arnedos, M. *et al.* Discordance between core needle biopsy (CNB) and excisional biopsy (EB) for estrogen receptor (ER), progesterone receptor (PgR) and HER2 status in early breast cancer (EBC). *Ann. Oncol.* **20,** 1948–1952 (2009).
23. Badoual, C. *et al.* Pathological prognostic factors of invasive breast carcinoma in ultrasound-guided large core biopsies-correlation with subsequent surgical excisions. *Breast* **14,** 22–27 (2005).
24. Cahill, R. A., Walsh, D., Landers, R. J. & Watson, R. G. Preoperative profiling of symptomatic breast cancer by diagnostic core biopsy. *Ann. Surg. Oncol.* **13,** 45–51 (2006).
25. Loubeyre, P. *et al.* Concordance between core needle biopsy and surgical excision specimens for tumour hormone receptor profiling according to the St. Gallen Classification, in clinical practice. *Breast J.* **19,** 605–610 (2011).
26. Jeselsohn, R. M. *et al.* Digital Quantification of Gene Expression in Sequential Breast Cancer Biopsies Reveals Activation of an Immune Response. *PLos One* **8,** e64225 (2013).
27. Wenger, R. H., Rolfs, A., Marti, H. H., Bauer, C. & Gassmann, M. Hypoxia, a novel inducer of acute phase gene expression in a human hepatoma cell line. *J. Biol. Chem.* **270,** 27865–27870 (1995).
28. Geyer, F. C. *et al.* Molecular analysis reveals a genetic basis for the phenotypic diversity of metaplastic breast carcinomas. *J. Pathol.* **220,** 562–573 (2010).
29. Du, P., Kibbe, W. A. & Lin, S. M. lumi: a pipeline for processing Illumina microarray. *Bioinformatics* **24,** 1547–1548 (2008).
30. Durinck, S. *et al.* BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* **21,** 3439–3440 (2005).

31. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4,** 1184–1191 (2009).
32. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8,** 118–127 (2007).
33. Haibe-Kains, B., Schroeder, M., Bontempi, G., Sotiriou, C. & Quackenbush, J. genefu: Relevant Functions for Gene Expression Analysis, Especially in Breast Cancer. *R package version* 1.8.0 (2012).
34. Schwender, H. siggenes: Multiple testing using SAM and Efron's empirical Bayes approaches. (2012).
35. Kitchen, R. R. *et al.* Correcting for intra-experiment variation in Illumina BeadChip data is necessary to generate robust gene-expression profiles. *BMC Genomics* **11,** 134 (2010).
36. Fowler, T., Sen, R. & Roy, A. L. Regulation of Primary Response Genes. *Mol. Cell* **44,** 348–360 (2011).
37. Desmedt, C. *et al.* Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes. *Clin. Cancer Res.* **14,** 5158–5165 (2008).
38. Chang, H. Y. *et al.* Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. *Proc. Natl. Acad. Sci. USA* **102,** 3738–3743 (2005).
39. Bland, K. I., Konstadoulakis, M. M., Vezeridis, M. P. & Wanebo, H. J. Oncogene protein co-expression. Value of Ha-ras, c-myc, c-fos, and p53 as prognostic discriminants for breast carcinoma. *Ann. Surg.* **221,** 706–718, discussion 718–20 (1995).
40. Ronski, K. *et al.* Early growth response gene 1 (EGR1) is deleted in estrogen receptor-negative human breast carcinoma. *Cancer* **104,** 925–930 (2005).
41. Langer, S. *et al*. Jun and Fos family protein expression in human breast cancer: correlation of protein expression and clinicopathological parameters. *Eur. J. Gynaecol. Oncol.* **27,** 345–352 (2006).
42. Vleugel, M. M., Greijer, A. E., Bos, R., van der Wall, E. & van Diest, P. J. c-Jun activation is associated with proliferation and angiogenesis in invasive breast cancer. *Hum. Pathol.* **37,** 668–674 (2006).
43. Wolford, C. C. *et al.* Transcription factor ATF3 links host adaptive response to breast cancer metastasis. *J. Clin. Invest.* **123,** 2893–2906 (2013).
44. Kataoka, F. *et al.* EGRI and FOSB gene expressions in cancer stroma are independent prognostic indicators for epithelial ovarian cancer receiving standard therapy. *Genes. Chromosomes Cancer* **51,** 300–312 (2012).
45. Huang, J. *et al.* Effects of ischemia on gene expression. *J. Surg. Res.* **99,** 222–227 (2001).
46. Dash, A. *et al.* Changes in Differential Gene Expression because of Warm Ischemia Time of Radical Prostatectomy Specimens. *Am. J. Pathol.* **161,** 1743–1748 (2002).
47. López-Knowles, E. *et al.* Heterogeneity in global gene expression profiles between biopsy specimens taken peri-surgically from primary ER-positive breast carcinomas. *Breast Cancer Res.* **18,** 39 (2016).
48. Borgan, E. *et al.* Ischemia caused by time to freezing induces systematic microRNA and mRNA responses in cancer tissue. *Mol. Oncol.* **5,** 564–576 (2011).
49. Aktas, B. *et al.* Global gene expression changes induced by prolonged cold ischemic stress and preservation method of breast cancer tissue. *Mol. Oncol.* **8,** 717–727 (2014).
50. Campbell, M. J. *et al.* Proliferating macrophages associated with high grade, hormone receptor negative breast cancer and poor clinical outcome. *Breast Cancer Res. Treat.* **128,** 703–711 (2011).
51. Morrogh, M. *et al.* Differentially Expressed Genes in Window Trials are Influenced by the Wound-Healing Process: Lessons Learned from a Pilot Study with Anastrozole. *J. Surg. Res.* **176,** 121–132 (2012).
52. Dunbier, A. K. *et al.* Molecular profiling of aromatase inhibitor-treated post-menopausal breast tumors identifies immune-related correlates of resistance. *Clin Cancer Res* **19,** 2775–2786 (2013).
53. Chang, H. Y. *et al.* Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds. *PLoS Biol.* **2,** E7 (2004).
54. DeMichele, A. *et al.* The Neoadjuvant Model Is Still the Future for Drug Development in Breast Cancer. *Clin. Cancer Res.* **21,** 2911–2915 (2015).
55. Ong, K. R. *et al.* Biomarkers of dietary energy restriction in women at increased risk of breast cancer. *Cancer Prev Res* (*Phila Pa*) **2,** 720–731 (2009).
56. Wilson, C. L., Sims, A. H., Howell, A., Miller, C. J. & Clarke, R. B. Effects of oestrogen on gene expression in epithelium and stroma of normal human breast tissue. *Endocr Relat Cancer* **13,** 617–628 (2006).
57. Smith, I. E. *et al.* A phase II placebo-controlled trial of neoadjuvant anastrozole alone or with gefitinib in early breast cancer. *J. Clin. Oncol.* **25,** 3816–3822 (2007).
58. Sabine, V. S. *et al.* Gene expression profiling of response to mTOR inhibitor everolimus in pre-operatively treated post-menopausal women with oestrogen receptor-positive breast cancer. *Breast Cancer Res. Treat.* **122,** 419–428 (2010).
59. Magbanua, M. J. M. *et al.* Serial expression analysis of breast tumors during neoadjuvant chemotherapy reveals changes in cell cycle and immune pathways associated with recurrence and response. *Breast Cancer Res.* **17,** 73 (2015).

Acknowledgements

We are grateful for funding from Breast Cancer Now and to CRUK for a PhD stipend for DAP.

Author Contributions

A.H.S., J.M.D. and J.M.S.B. conceived the study and designed the experiments. L.R., J.S.T. and J.M.D. collected and reviewed samples. V.S.S., L.M.A. and A.K.T. carried out experiments. D.A.P., L.M.A., A.K.T. and A.H.S. performed the analysis. D.A.P. and A.H.S. drafted the paper. All authors approved the final manuscript.

Additional Information

Accession Code: The raw and processed data from this study can be accessed from NCBI GEO under the accession GSE76728.

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Pearce, D. A. *et al.* Tumour sampling method can significantly influence gene expression profiles derived from neoadjuvant window studies. *Sci. Rep.* **6**, 29434; doi: 10.1038/srep29434 (2016).

Continuous Biomarker Assessment by Exhaustive Survival Analysis

Dominic A. Pearce[1], Ajit J. Nirmal[2], Tom C. Freeman[2], Andrew H. Sims[1*]

[1]Applied Bioinformatics of Cancer, University of Edinburgh Cancer Research Centre, Institute of Genetics and Molecular Medicine, Edinburgh, UK [2]Systems Immunology Group, Division of Genetics and Genomics, The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian, EH25 9RG

*Corresponding author: andrew.sims@ed.ac.uk

ABSTRACT

Publicly available high-throughput molecular data can enable biomarker identification and evaluation in a meta-analysis. However, a continuous biomarker's underlying distribution and/or potential confounding factors associated with outcome will inevitably vary between cohorts and is often ignored. The *survivALL* R package (https://CRAN.R-project.org/package=survivALL) allows researchers to generate visual and numerical comparisons of all possible points-of-separation, enabling quantitative biomarkers to be reliably evaluated within and across datasets, independent of compositional variation. Here, we demonstrate *survivALL's* ability to robustly and reproducibly determine an applicable level of gene expression for patient prognostic classification, in datasets of similar and dissimilar compositions. We believe *survivALL* represents a significant improvement over existing methodologies in stratifying patients and determining quantitative biomarker(s) cut-points for public and novel datasets.

INTRODUCTION

Biomarker performance is traditionally assessed by estimating survival benefit over time, often summarized by Kaplan-Meier plot[1,2]. Whilst qualitative biomarkers can easily stratify cohorts into two or more groups, the question of how to divide a cohort with a quantitative measure is considerably more challenging. Multiple subgroups may exist within a given patient population, but a need for simple and tractable clinical decisions (i.e. treat or don't treat) have often encouraged division into two classes. The most obvious, common, yet arbitrary approach for this is to divide a cohort into two equal sized groups at the median level. However, this median-split approach ignores a marker's distribution and any potential confounding factors relating to the composition of the dataset, clinical or otherwise. This heterogeneity makes assessing the robustness of a biomarker across datasets problematic, with median dichotomisation unlikely to produce reliable results due to random sampling differences. In short, a significant median separation in cohort A is unlikely to translate to cohort B. Given the public availability of high-throughput molecular profiling datasets providing outcome data, the opportunity to evaluate biomarkers *in silico,* as part of extended meta-analyses, is in principle becoming easier. In practice however, datasets derived from cohorts of primary patient tissue are molecularly heterogeneous in their composition[3], are likely to exhibit varying proportions of multiple clinical factors associated with outcome (e.g. node and receptor statuses, grade, molecular subtypes and age). It is therefore clear that this variation must be accounted for if a biomarker is to be successfully applied and validated in a meta-analysis.

bioRxiv preprint first posted online Oct. 25, 2017; doi: http://dx.doi.org/10.1101/208660. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Building upon previous best-of-split methods[4–6] we have developed *survivALL* (https://CRAN.R-project.org/package=survivALL) an R package implementation to exhaustively calculate and visualise hazard ratios (HRs) for all possible points-of-separation and assess the association between a continuous measure and survival. As open-source software, *survivALL* allows for researchers to reproducibly perform exhaustive survival analysis using public or independent datasets, in a highly transparent, automatable and extensible manner within the wider R package landscape. Complimenting *survivALL* is a companion web-based app – *survivAPP* (pearcedom.shinyapps.io/survivapp/) – allowing non-programmatic, drag-and-drop *survivALL* use. We believe that *survivALL* allows researchers to perform exhaustive survival analysis relevant to the current state of –omics research.

MATERIALS AND METHODS

survivALL computes hazard ratio statistics for every point-of-separation possible, allowing the magnitude, and frequency of, significant cut-points to be identified (Video S1)x. To more robustly determine significance, a non-parametric and dataset-specific bootstrap is applied, defining reliable confidence intervals. Bootstrapping is performed as a 10,000-fold repeated calculation of HRs for random sample orderings, producing a distribution of expected/random HRs for each individual point-of-separation, to which observed true biological HRs are compared and significance calculated. Dataset-specific best points-of-separation were optimised by maximising desirability of combined p-value and absolute HR magnitude[7].

All analysis was performed using the R statistical environment using CRAN and Bioconductor hosted packages[8–10]. *survivALL* p-values were calculated using the bootstrapping procedure described above, whilst median approach p-values were calculated using the Cox-proportional hazards model[1]. To avoid overcorrection of the multiple comparisons inherent in *survivALL*, all *survivALL* analysis is preceded by a single ancillary test of significance (Cox-proportional hazards model), performed using the biomarker as a continuous variable. Code and session information necessary to reproduce the analysis is included as supplementary material.

RESULTS

To illustrate the value of the *survivALL* package, we considered over-expression of human epidermal growth factor 2 (HER2/*ERBB2*), a well-established biomarker associated with poor prognosis in invasive breast cancer[11,12], in the largest single breast cancer gene expression dataset – METABRIC[13] (*n=1971* samples with complete disease specific survival information). Importantly and unusually, METABRIC is split into two equally sized and composition-matched subsets, allowing for independent discovery and validation. The *survivALL* plotALL() function enables visualisation of hazard ratios for all possible separations (n-1) of ordered *ERBB2* expression for the METABRIC discovery cohort (*n=980*) (Figure 1A). It establishes 283/980 points-of-separation as significant, confirming the established expectation that increased *ERBB2* expression is associated with poor prognosis. The span and location of these points are consistent with epidemiological evidence demonstrating that across the population ~20% of breast cancers overexpress HER2[14]. Contrasted in Figure 1B are two Kaplan Meier plots resulting from using the median (p=0.48) and the data-driven most significant point-of-separation ($p=1.25x10^{-11}$) for this dataset.

In considering the issue of dataset composition, the highly similar METABRIC subsets enabled us to first evaluate whether *survivALL* could derive and validate the point-of-separation which

most clearly distinguishes (lowest p-value) between good and poor prognosis groups from one cohort to the other (Figure 1B). For all 19,628 genes, *survivALL* was used to calculate the point-of-separation with the lowest p-value in the discovery cohort and the gene expression value at this point was used to divide the validation cohort (*n=991*). Prediction accuracy was measured as the number of patients incorrectly classified compared to the validation cohort's own dataset-specific most significant point-of-separation. This same measure of accuracy was additionally calculated using a median approach for comparison (Figure 1B & 1C).

As expected, *survivALL* significantly outperforms a median approach ($p=2.2x10^{-16}$, Kolmogorov-Smirnov test), demonstrating *survivALL*'s applicability in robustly determining prognostic stratification in two suitably relatable datasets, with no *a priori* knowledge of a gene's population level distribution.

However, real-world datasets are rarely compositionally matched in the way that the METABRIC discovery and validation cohorts are, and it remained to determine if *survivALL* could perform this applied prognostic stratification in more dissimilar and noisier datasets. To test this, we simulated semi-random sub-samplings of the entire METABRIC dataset for three pre-defined proportions of estrogen receptor (ER) positivity – 25, 50 & 75%. Using these dramatically variable compositions, we determined the extent to which *survivALL* was able to track these differences in terms of the datasets-specific best point-of-separation of *GATA3*, a mediator of ER binding[15] (Figure 2A). It was evident that as the proportion of ER+ and ER- samples shifted, the *survivALL* plots in turn shifted in response, with the most significant point-of-separation consistently falling at the division between ER+ and ER- samples. Importantly, though these plots changed, the corresponding level of expression that defined the ER+/- boundary remained consistent (Figure 2B), indicating *survivALL's* ability to robustly determine a reliable level of GATA3 expression to stratify our patient cohorts, even in highly compositionally dissimilar datasets. We repeated this analysis for 5 publically available, ER-positive and tamoxifen treated datasets (GSE2990, GSE6532, GSE9195, GSE12093, GSE17705), for grade and AURKA expression, demonstrating similarly consistent results between datasets (Figure S1).

Finally, whilst significant association of a biomarker can be determined in a meta-analysis using expression as a continuous variable, this does not reveal the direction of that association, i.e. good or poor prognosis. Beyond the added information revealed therein, there additionally remains the possibility that a gene determined as significantly associated with prognosis in more than one dataset may in fact demonstrate variable, or even opposite, directions of association (Figure S2). For the METABRIC discovery and validation cohorts this was evident to occur for ten genes, several of which (*ACY3*[16], *LRRK2*[17], *NUPR1*[18] & *UGT1A7*[19]) have been previously associated with cancer risk. This therefore represents a small but real danger that must be considered in meta-analysis.

DISCUSSION

In this study we have highlighted the issue of assessing quantitative biomarkers for survival analysis, offering *survivALL* and *survivAPP* (Video S2) as tools to evaluate and overcome these challenges. We have attempted to illustrate situations relevant and common to researchers in the current environment of publically available large datasets, using well established but likely compositionally distorted examples. As an R package, *survivALL* allows greater resolution, transparency and flexibility compared to other online best-of-split tools, applicable to any public or proprietary dataset and usable in larger-scale automated data operations. Researchers are increasingly using approaches such as KMplotter[20], either to median-split or highlight the

bioRxiv preprint first posted online Oct. 25, 2017; doi: http://dx.doi.org/10.1101/208660. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

dataset-specific best point-of-separation in a single or combined dataset. This approach presents a number of potential problems, most notably the restriction of what data is available to be analysed – either alone or in combination as a meta-analysis – and the exact methods used. Furthermore, whilst a prognostic association can be considered using a continuous marker as the classifier itself, this ignores a number of informative factors, such as the direction and magnitude of association, as well as the optimal value to allow patient stratification into treatment groups.

Importantly, whilst the examples presented here relate to individual genes in breast cancer microarray datasets, *survivALL* is readily extensible to other diseases, data types or any other quantitative measure, including signatures or scores based upon combinations of markers. Moreover, whilst this paper has largely considered dichotomisation, *survivALL* also allows for additional sub-populations with varying survivals to be visualised, including those related to clinical factors such as grade (Figure S1), as well as revealing potential confounding factors, producing multivariate analysis and demonstrating or uncovering an otherwise unknown factor when considering multiple datasets. Fundamentally, *survivALL* has been developed as open-source software, to flexibly integrate with other popular R packages, including the popular visualisation tool *ggplot2* for customisable and scalable output.

Most importantly, *survivALL* and *survivAPP* allows true biological effects and their relationship to survival to be revealed and reliably compared, within and between datasets, to move towards determining real-world clinically applicable biomarker cut-offs.

References

1. Cox, D. R. Regression Models and Life-Tables. *J. R. Stat. Soc. Ser. B* **34,** 187–220 (1972).

2. Kaplan, E. L. & Meier, P. Nonparametric Estimation from Incomplete Observations. *J. Am. Stat. Assoc.* **53,** 457–481 (1958).

3. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406,** 747–52 (2000).

4. Camp, R. L., Dolled-Filhart, M. & Rimm, D. L. X-Tile: A New Bio-Informatics Tool for Biomarker Assessment and Outcome-Based Cut-Point Optimization. *Clin. Cancer Res.* **10,** 7252–7259 (2004).

5. Motakis, E., Ivshina, A. & Kuznetsov, V. Data-driven approach to predict survival of cancer patients. *IEEE Eng. Med. Biol. Mag.* **28,** 58–66 (2009).

6. Mizuno, H., Kitada, K., Nakai, K. & Sarai, A. PrognoScan: a new database for meta-analysis of the prognostic value of genes. *BMC Med. Genomics* **2,** 18 (2009).

7. Lazic, S. E. Ranking, selecting, and prioritising genes with desirability functions. *PeerJ* **3,** e1444 (2015).

8. Therneau, T. M. Survival Analysis [R package survival version 2.41-3]. (2015).

9. Schröder, M. S., Culhane, A. C., Quackenbush, J. & Haibe-Kains, B. survcomp: an R/Bioconductor package for performance assessment and comparison of survival models. *Bioinformatics* **27,** 3206–3208 (2011).

10. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis.* (Springer-Verlag New York, 2009).

11. Kapp, A. V *et al.* Discovery and validation of breast cancer subtypes. *BMC Genomics* **7,** 231 (2006).

12. Mackay, A. *et al.* Microarray-based class discovery for molecular classification of breast cancer: analysis of interobserver agreement. *J. Natl. Cancer Inst.* **103,** 662–73 (2011).

13. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486,** 346–52 (2012).

14. Dean-Colomb, W. & Esteva, F. J. Her2-positive breast cancer: Herceptin and beyond. *Eur. J. Cancer* **44,** 2806–2812 (2008).

15. Theodorou, V., Stark, R., Menon, S. & Carroll, J. S. GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility. *Genome Res.* **23,** 12–22 (2013).

16. Long, P. M., Stradecki, H. M., Minturn, J. E., Wesley, U. V. & Jaworski, D. M. Differential aminoacylase expression in neuroblastoma. *Int. J. Cancer* **129,** 1322–1330 (2011).

17. Agalliu, I. *et al.* Higher Frequency of Certain Cancers in *LRRK2* G2019S Mutation Carriers With Parkinson Disease. *JAMA Neurol.* **72,** 58 (2015).

18. Emma, M. R. *et al.* NUPR1, a new target in liver cancer: implication in controlling cell growth, migration, invasion and sorafenib resistance. *Cell Death Dis.* **7,** e2269–e2269 (2016).

19. Carlini, L. E. *et al.* UGT1A7 and UGT1A9 polymorphisms predict response and toxicity in colorectal cancer patients treated with capecitabine/irinotecan. *Clin. Cancer Res.* **11,** 1226–36 (2005).

20. Szász, A. M. *et al.* Cross-validation of survival associated biomarkers in gastric cancer using transcriptomic data of 1,065 patients. *Oncotarget* (2014). doi:10.18632/oncotarget.10337

bioRxiv preprint first posted online Oct. 25, 2017; doi: http://dx.doi.org/10.1101/208660. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

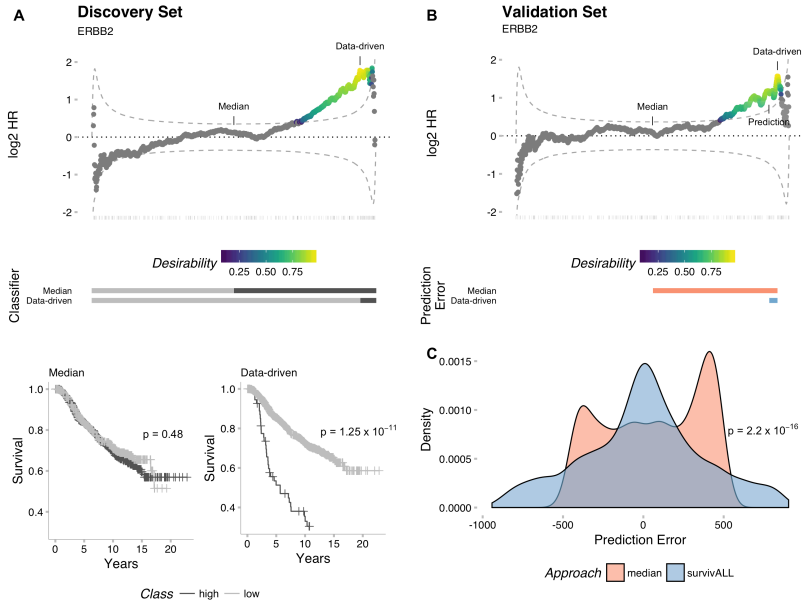

Figure 1: A) The plotALL function allows hazard ratios (y-axis) and bootstrap significance ($p<0.05$ colour scale; p 0.05 grey) for all possible quantitative biomarker cut-points to be examined. Patients are ordered by increasing expression of the quantitative marker – in this example, *ERBB2* expression. Events (distant metastasis-free survival) are indicated as vertical lines below the plot, with darker colours indicating chronologically earlier events. Kaplan-Meier plots for *ERBB2* stratified using median (n = 490 vs. 490) and survivALL (n = 924 vs. 56) approaches are shown, comparing patients with expression lower (grey) or higher (black) than the points-of-separation determined in A. B) Comparison of *survivALL* and median approaches in predicting optimal cohort separation for *ERBB2*. The expression values associated with both median and dataset-specific best points-of-separation derived from the discovery set in 1A is applied to the validation dataset. Error (median = orange, *survivALL* = blue) is calculated as the number of patient who would be incorrectly assigned using these predictions, in relation to the validation cohorts own best separation. This process is repeated in 1C for all genes. *survivALL* significantly ($p=2.2x10^{-16}$) outperforms a median approach in terms of predictive accuracy.

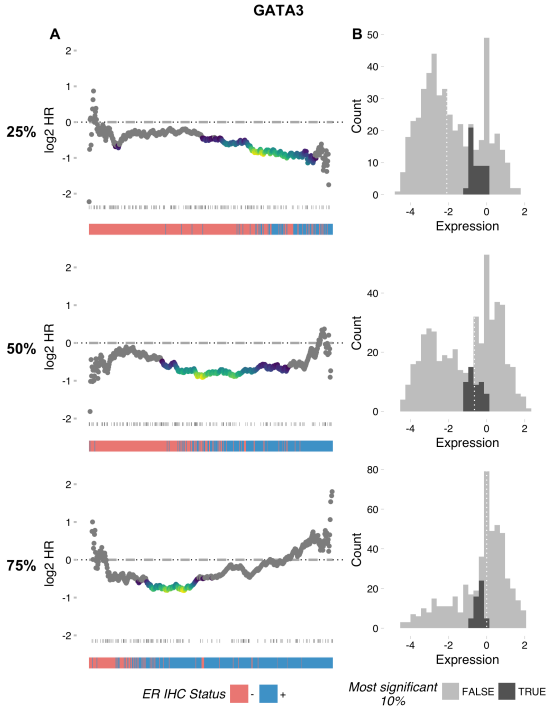

Figure 2: A) Analysis of compositionally *un*matched METABRIC subsets. Three equally sized subsets (n=500) with variable proportions of ER+ samples (25, 50 and 75%) demonstrate *survivALL*'s ability to track the ER+/ER- (blue/red) boundary for *GATA3* expression, with the most significant point-of-separation (low-high significance = blue-yellow gradient) existing at this boundary for each subset. B) Whilst the HR distribution is evident to change with ER+ proportion, the 10% most significant points-of-separation consistently related to the same level of *GATA3* expression. For comparison, the median level of expression (dashed white line) for each subset is also shown.

Breast Cancer Research

RESEARCH ARTICLE Open Access

Molecular changes during extended neoadjuvant letrozole treatment of breast cancer: distinguishing acquired resistance from dormant tumours

Cigdem Selli[1,2], Arran K. Turnbull[1,3], Dominic A. Pearce[1], Ang Li[1], Anu Fernando[1,3], Jimi Wills[4], Lorna Renshaw[3], Jeremy S. Thomas[3], J. Michael Dixon[3] and Andrew H. Sims[1*]

Abstract

Background: The risk of recurrence for endocrine-treated breast cancer patients persists for many years or even decades following surgery and apparently successful adjuvant therapy. This period of dormancy and acquired resistance is inherently difficult to investigate; previous efforts have been limited to in-vitro or in-vivo approaches. In this study, sequential tumour samples from patients receiving extended neoadjuvant aromatase inhibitor therapy were characterised as a novel clinical model.

Methods: Consecutive tumour samples from 62 patients undergoing extended (4–45 months) neoadjuvant aromatase inhibitor therapy with letrozole were subjected to transcriptomic and proteomic analysis, representing before (≤ 0), early (13–120 days), and long-term (> 120 days) neoadjuvant aromatase inhibitor therapy with letrozole. Patients with at least a 40% initial reduction in tumour size by 4 months of treatment were included. Of these, 42 patients with no subsequent progression were classified as "dormant", and the remaining 20 patients as "acquired resistant".

Results: Changes in gene expression in dormant tumours begin early and become more pronounced at later time points. Therapy-induced changes in resistant tumours were common features of treatment, rather than being specific to the resistant phenotype. Comparative analysis of long-term treated dormant and resistant tumours highlighted changes in epigenetics pathways including DNA methylation and histone acetylation. The DNA methylation marks 5-methylcytosine and 5-hydroxymethylcytosine were significantly reduced in resistant tumours compared with dormant tissues after extended letrozole treatment.

Conclusions: This is the first patient-matched gene expression study investigating long-term aromatase inhibitor-induced dormancy and acquired resistance in breast cancer. Dormant tumours continue to change during treatment whereas acquired resistant tumours more closely resemble their diagnostic samples. Global loss of DNA methylation was observed in resistant tumours under extended treatment. Epigenetic alterations may lead to escape from dormancy and drive acquired resistance in a subset of patients, supporting a potential role for therapy targeted at these epigenetic alterations in the management of resistance to oestrogen deprivation therapy.

Keywords: Dormancy, Oestrogen deprivation therapy, Epigenetics, Letrozole, Sequential samples, Resistance, Microarray, Proteomics

* Correspondence: andrew.sims@ed.ac.uk
[1]Applied Bioinformatics of Cancer, University of Edinburgh Cancer Research UK Centre, MRC Institute of Genetics and Molecular Medicine, Edinburgh, UK
Full list of author information is available at the end of the article

Background

Approximately 70% of breast cancer patients who have oestrogen receptor (ER) alpha-positive tumours receive adjuvant oestrogen deprivation therapy. Five years of aromatase inhibitor therapy produces a 40% reduction in 10-year mortality [1]. However, while the annual risk of mortality for ER-negative breast cancer decreases following the first 5 years after diagnosis, the annual rate remains constant for ER^+ patients [2]. In fact, women with ER^+ early-stage disease treated with 5 years of adjuvant endocrine therapy have a persistent risk of recurrence and death from breast cancer for at least 20 years after diagnosis [3]. Molecular studies have demonstrated that nodal and distant metastases are highly similar to their matched primary tumours, implicating a continuation of the original cancer [4–6]. However, the time between treatment and recurrence is often greater than that which can be explained by normal cell-doubling rates [7], implying cancer cells remain dormant in the body before re-awakening.

Residual dormant cancer cells are hypothesised to persist either by withdrawing from the cell cycle and transitioning to a quiescence state or by continuing to proliferate at a reduced rate, counter-balanced by cell death [8]. Reawakened dormant cells may become detectable after reaching a detection threshold or reactivated via increased angiogenesis, and/or escape from the inhibitory microenvironment or immune effects [9, 10]. Dormancy is therefore considered a major mechanism underlying resistance to therapy, where dormant cells survive despite anti-proliferative oestrogen deprivation therapy.

Resistance to oestrogen deprivation therapy may occur at disease inception (de novo or innate resistance), but a larger proportion of patients acquire resistance during treatment (acquired/secondary resistance) [11]. Several mechanisms of resistance to oestrogen deprivation therapy have been described previously [12, 13]. However, the majority of these findings are based on preclinical data obtained from cell lines and animal models. It is therefore difficult to know if these accurately reflect molecular changes in patient tumours.

Expression profiling of clinical samples, measuring the effect of, or predicting response to, treatment has recently become feasible. However, experimental design issues, such as the difficulty in obtaining paired samples for comparison, particularly for longer time intervals, makes it difficult to study changes within tumours [14]. For example, a previous study investigating tamoxifen failure compared samples from patients requiring salvage surgery with pre-treatment samples from an unrelated group of disease-free patients [15]. More recently, sequential patient-matched samples have been successfully utilised to determine treatment-induced dynamic changes in tumours at 2 weeks to 3 months, demonstrating the effectiveness of this approach [16–18].

For a variety of reasons, including being unfit for surgery, a proportion of patients receiving pre-surgical oestrogen deprivation therapy do not have their tumours excised following 3–4 months of treatment. These long-term endocrine-treated tumours represent a unique group that can inform how tumours respond to extended oestrogen deprivation in situ. Having initially shrunk in size, some tumours remain at a steady volume and appear dormant, whilst others subsequently begin to regrow. We have utilised this unique cohort of sequential samples from patients receiving extended neoadjuvant oestrogen deprivation therapy to characterise luminal breast cancer dormancy and acquired resistance as a novel clinical model.

Methods

Patients and samples

Breast cancer patients were treated with neoadjuvant aromatase inhibitor therapy with letrozole (Femara, 2.5 mg; Novartis Pharma AG, Basel, Switzerland) for a minimum of 4 months; tumours were not removed either because patients declined or were unfit for surgery. The study was approved by the local regional ethics committee (07/S1103/26, August 2007) and all patients gave informed consent. Clinical characteristics of the tumours are given in Table 1. A consort diagram detailing inclusion and exclusion criteria is provided in Additional file 1 (Figure S1). Patients with > 40% initial decrease in tumour size by 4 months of treatment were included in the study. Those with no subsequent progression on imaging by the latest biopsy were classified as "dormant"; otherwise, they were classified as "acquired resistant" (Fig. 1a, b). For patients whose latest ultrasound scan (USS) measurement was taken more than a month before surgery, changes in gene expression (mean relative change) of three widely used proliferation markers (MKI67, PCNA, and MCM2) were used to assist classification. Tumours with an increase in proliferation marker expression (either after an initial decrease or not) were classified as "acquired resistant", otherwise there were classified as "dormant". Sequential tumour biopsies were taken with a 14-gauge needle before and after letrozole treatment and at the time of surgery. Fresh samples were snap-frozen in liquid nitrogen and each tumour sample was confirmed to contain ≥ 50% cellularity and at least 60% tumour tissue using haematoxylin and eosin (H&E) sections. Following pulverisation of tissue with a membrane disruptor (Micro-Dismembrator U, Braun Biotech), phase separation was performed by guanidinium thiocyanate-phenol-chloroform extraction (Qiazol Lysis Reagent).

Table 1 Patient characteristics

	Dormant, *n* (%)	Resistant, *n* (%)	Total, *n*	*p* value[a]
Total no. of patients	42	20	62	
Total no. of samples	111	56	167	
Age at diagnosis				
Mean	75	72		
Range	53–87	56–89		
Tumour grade				0.39
1	6 (14.3)	1 (5.0)	7	
2	27 (64.3)	10 (50.0)	37	
3	8 (19.0)	6 (30.0)	14	
NA	1 (2.4)	3 (15.0)	4	
Tumour size				0.71
T1	5 (11.9)	4 (20.0)	9	
T2	19 (45.2)	9 (45.0)	28	
T3	2 (4.8)	2 (10.0)	4	
T4	11 (26.2)	4 (20.0)	15	
NA	5 (11.9)	1 (5.0)	6	
Nodal status				0.36
N0	27 (64.3)	11 (55.0)	38	
N1	8 (19.0)	7 (35.0)	15	
N2	1 (2.4)	1 (5.0)	2	
N3	1 (2.4)	0	1	
NX	1 (2.4)	0	1	
NA	4 (9.5)	1 (5.0)	5	
Metastasis status				1.00
M0	34 (80.9)	18 (90.0)	56	
M1	2 (4.8)	0	2	
MX	1 (2.4)	1 (5.0)	2	
NA	5 (11.9)	1 (5.0)	6	
ER score (Allred)				0.18
6	1 (2.4)	0	1	
7	6 (14.3)	6 (30.0)	12	
8	35 (83.3)	14 (70.0)	49	
HER status				0.69
Negative	35 (83.3)	12 (60.0)	47	
Positive	6 (14.3)	3 (15.0)	9	
NA	1 (2.4)	5 (25.0)	6	
Histological type				0.73
IDC (no special type)	18 (42.9)	6 (30)	24	
ILC	8 (19.0)	4 (20)	12	
Mucinous	1 (2.4)	0	1	
NA	15 (35.7)	10 (50)	25	
Molecular subtype[b]				1.00
Luminal A	21 (50.0)	9 (45.0)	30	
Luminal B	20 (47.6)	9 (45.0)	29	

Table 1 Patient characteristics *(Continued)*

	Dormant, *n* (%)	Resistant, *n* (%)	Total, *n*	*p* value[a]
HER2 enriched	0	1 (5.0)	1	
Basal-like	0	0	0	
Normal-like	0	0	0	
NA	1 (2.4)	1 (5.0)	2	

ER oestrogen receptor, *IDC* invasive ductal carcinoma, *ILC* invasive lobular carcinoma, *NA* not available
[a]Fisher exact test ($p < 0.05$, two-tailed)
[b]At diagnosis by PAM50 (genefu)

Gene expression profiling and analysis

RNA was extracted from the aqueous phase by column-based purification (RNeasy mini kit, Qiagen) and then labelled and hybridized (HumanHT-12 v4 Illumina BeadChip) according to the manufacturer's protocol (NuGEN) as previously described [19, 20]. Raw data were detection ($p < 0.05$, ≥ 3 samples) and quality filtered, $\log_2$ transformed, and quantile normalized using the Bioconductor lumi package [21]. Data are publicly available from NCBI GEO under accession GSE111563. The analysis also includes data from 14 patients (42 samples, GSE59515) and 9 patients (24 samples, GSE55374) from previous studies [16, 19] which meet the criteria defined above; the relationship between the samples from these datasets is indicated in Additional file 2 (Table S1). Hierarchical clustering analysis was performed using a complete linkage method and Euclidean distance. Pathway enrichment analysis and visualisation were performed using ReactomePA [22]. Differential gene expression analysis was performed with Rank Products [23]. The significance of differences was evaluated by using unpaired Wilcoxon test for two groups and analysis of variance (ANOVA) with post-hoc Tukey HSD for multiple comparisons.

Proteomics analysis

Proteins were isolated from the organic phase of Qiazol [24]. Pellets were sonicated and dissolved in 1% SDS. Proteomics was performed using Thermo Q Exactive plus and Label-free Quantitation (LFQ). Peptides obtained from samples were analysed in mass spectrometry runs; serial samples from the patients were run on the same day. A modified version of Filter Aided Sample Preparation (FASP) was performed using serial digests with lysC and trypsin to generate two orthogonal fractions per sample [25, 26]. The mass spectrometry spectra generated in each run were used for relative quantitation of individual peptides. Normalization and quantifications of peptides were performed using MaxLFQ and MaxQuant [27]. A total of 6251 protein groups were identified. Data was $\log_2$ transformed and missing values were imputed as the minimum observed value in each sample. The data have been deposited to the ProteomeXchange Consortium via the PRIDE [28] partner repository with the dataset identifier PXD009328.

Immunohistochemistry and scoring

Formalin-fixed paraffin-embedded (FFPE) sections were processed using an automated stainer (Leica Biosystems, Bond III). Heat-induced epitope retrieval for both antibodies was performed by 30-min incubation in citrate-based pH 6.0 epitope retrieval (ER1) solution followed by incubation in 3.5 N HCl for 15 min at room temperature as suggested by Haffner et al. [29]. For 5-methylcytosine (5-mC) and 5-hydroxymethylcytosine (5-hmC) detection, mouse monoclonal 5-methylcytosine specific (33D3; Abcam, ab10805) and rabbit polyclonal 5-hydroxylmethylcytosine (Active Motif, 39,769) antibodies were used, respectively. Both antibodies were used at 1/1000 dilution and were incubated for 15 min. Detection was performed using secondary antibody-horseradish peroxidase (HRP) conjugates and substrate-chromogen (3,3'-diaminobenzidine (DAB)). After staining, slides were counterstained with haematoxylin. Nuclear staining in epithelial cells was evaluated using an H-score obtained by multiplying the intensity of the stain (0: no staining; 1: weak staining; 2: moderate staining; 3: intense staining) by the percentage of cells (H-score range, 0 to 300).

Results

Long-term oestrogen deprivation therapy as a model of dormancy and acquired resistance

A cohort of 62 primary breast cancer patients receiving at least 4 months of oestrogen deprivation therapy (Fig. 1a) and initially responding were stratified into two groups, 'dormant' and 'acquired resistant' based on dynamic changes in tumour size and proliferation (see methods and Fig. 1b). Patient-matched sequential samples were available at three time points: before (≤ 0 days), early (13–120 days), and long-term (> 120 days) treatment. Dormant and acquired resistant samples were distributed uniformly with respect to time on treatment, and duration at each time point was not significantly different between response groups (Table 1). For long-term treatment, the mean and range were 186 (121–884) days

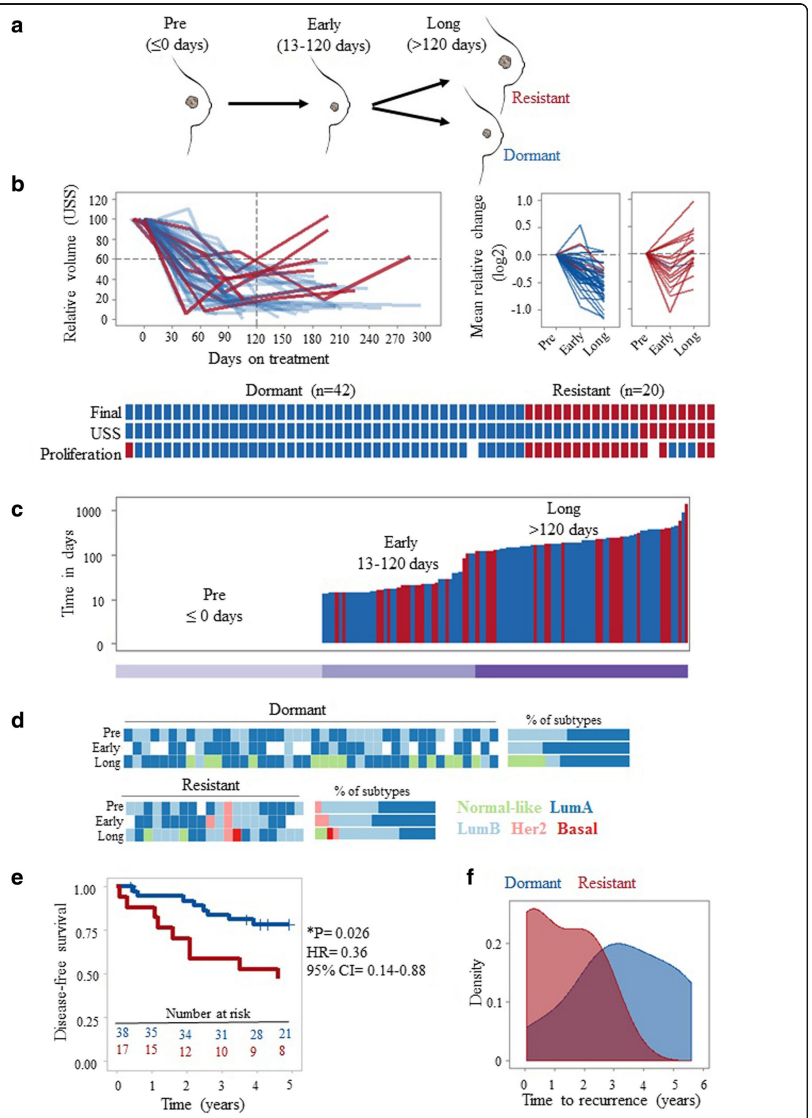

Fig. 1 (See legend on next page.)

(See figure on previous page.)

Fig. 1 Long-term oestrogen deprivation therapy as a clinical model to investigate breast cancer dormancy and acquired resistance. **a** Extended (4–45 months) letrozole treatment was exploited as a clinical model of breast cancer dormancy and acquired resistance. Sequential clinical samples from the same patient with no surgery and extended treatment were used to model clinical breast cancer dormancy and resistance. Before (pre, ≤ 0 days), early-on (early, 13–120 days) and long-term (long, > 120 days) neoadjuvant aromatase inhibitor therapy with letrozole. **b** Dynamic change in tumour size by ultrasound scan (USS) and mean expression of proliferation markers MKI67, PCNA, and MCM2 were used to classify patients into two categories: dormant (blue) and resistant (red). Overall comparisons of classifications per patient based on USS and mean change in proliferation markers with final classification are shown. **c** The duration of letrozole treatment (days) for samples from dormant (blue) and resistant (red) patients. Each bar represents a sample. Samples are ordered by time on treatment. **d** Intrinsic subtype classification by PAM50 of samples at each time point. Stacked bar graphs on the right show the percentage of each subtype of samples from dormant and resistant patients. **e** Kaplan-Meier plot showing disease-free survival probability in patients with dormant versus resistant tumours (log-rank test). Disease-free survival was defined from time of surgery. **f** Density plot showing the distribution of time to recurrence (in years; defined from time of surgery) in patients with dormant and resistant tumours. CI confidence interval, HR hazard ratio, LumA luminal A, LumB luminal B

and 226 (121–1366) days for patients with dormant and acquired resistant tumours, respectively (Fig. 1c).

There were no significant differences in patient clinico-pathological features between response classes before treatment (Table 1). However, prediction analysis of microarray (PAM)50 intrinsic molecular subtypes were found to change during oestrogen deprivation therapy (Fig. 1d). These changes were consistent with known associations with outcome, with all dormant tumours either remaining the same or switching to better prognosis luminal A or normal-like tumours. For resistant tumours, however, 25% (5 out of 20) switched to a subtype of worse prognosis (Fig. 1d). The proportion of luminal B tumours characterised by reduced endocrine sensitivity and higher proliferation was higher in resistant tumours compared with dormant tumours under early (35% versus 27%) and long-term (50% versus 12%) treatment (Fig. 1d; stacked bar graphs on the right). The PAM50 defines breast cancer into four intrinsic molecular subtypes: luminal A, luminal B, HER2-enriched, and basal-like [30]. PAM50 intrinsic subtyping has been shown to provide additional prognostic value to standard clinicopathological factors where luminal A tumours had a significantly better outcome than luminal B, HER2-enriched, and basal-like tumours [31].

As expected, Kaplan-Meier survival analysis demonstrated significantly worse outcomes for patients with resistant tumours compared with patients with dormant tumours (log rank, $p = 0.026$; Fig. 1e). Recurrence rates for patients with dormant and resistant tumours were 21% (9/42) and 45% (9/20), respectively. Moreover, patients with resistant tumours suffered significantly earlier recurrences compared with patients with dormant tumours ($p = 0.05$; range 26–947 versus 136–2042 days; Fig. 1f). Disease-free survival and time to recurrence were defined from time of surgery, not from the time of diagnosis, since patient classification was performed based on change in tumour size by USS and proliferation by gene expression at on-treatment and surgery time points.

Distinct transcriptomic changes under long-term letrozole treatment

Unsupervised analysis was performed to consider whether sequential samples displayed greater similarity between response classes or treatment duration. Hierarchical clustering using the 500 genes with the highest variance across all samples revealed two main subclasses, seemingly driven by time on treatment, with resistant and dormant tumours indistinguishable (Fig. 2a). The dominant pattern was that the samples of the same patient usually clustered together (Fig. 2a).

When long-term treated samples were considered alone, two clusters did emerge, the larger of which contained mostly dormant samples (79%), whilst the second had a roughly even proportion of dormant (48%) and resistant (52%) samples (Fig. 2a). Similarly, a multidimensional scaling (MDS) plot for the 500 genes with the highest variance across all time points revealed consistent changes over time in response to treatment for both dormant and acquired resistant samples (Fig. 2b), although long-term dormant samples were much more distinguishable from pre-treatment samples than the long-term acquired resistant samples (Fig. 2b).

Correlations between tumours from different individuals (inter-patient) remained similar at each time point and were not different between response classes (data not shown) as corroborated by hierarchical clustering analysis with all samples and across all the time points. However, correlations of the transcriptome between matched sequential samples (intra-patient) revealed that pre-treatment samples were significantly ($p = 0.01$) less similar to their long-term treated pairs (median 0.89, range 0.74–0.95) than their early treatment pairs (median 0.91, range 0.84–0.95) (Fig. 2c). However, when divided by dormancy status this finding was only significant ($p = 0.01$) for dormant tumours (Fig. 2c), suggesting that dormant tumours continue to diverge transcriptionally whereas acquired resistant tumours do not consistently differ after initial or extended treatment, as mirrored in the MDS representation (Fig. 2b).

Perhaps surprisingly, oestrogen, progesterone, and androgen receptors and their target genes [32] were not

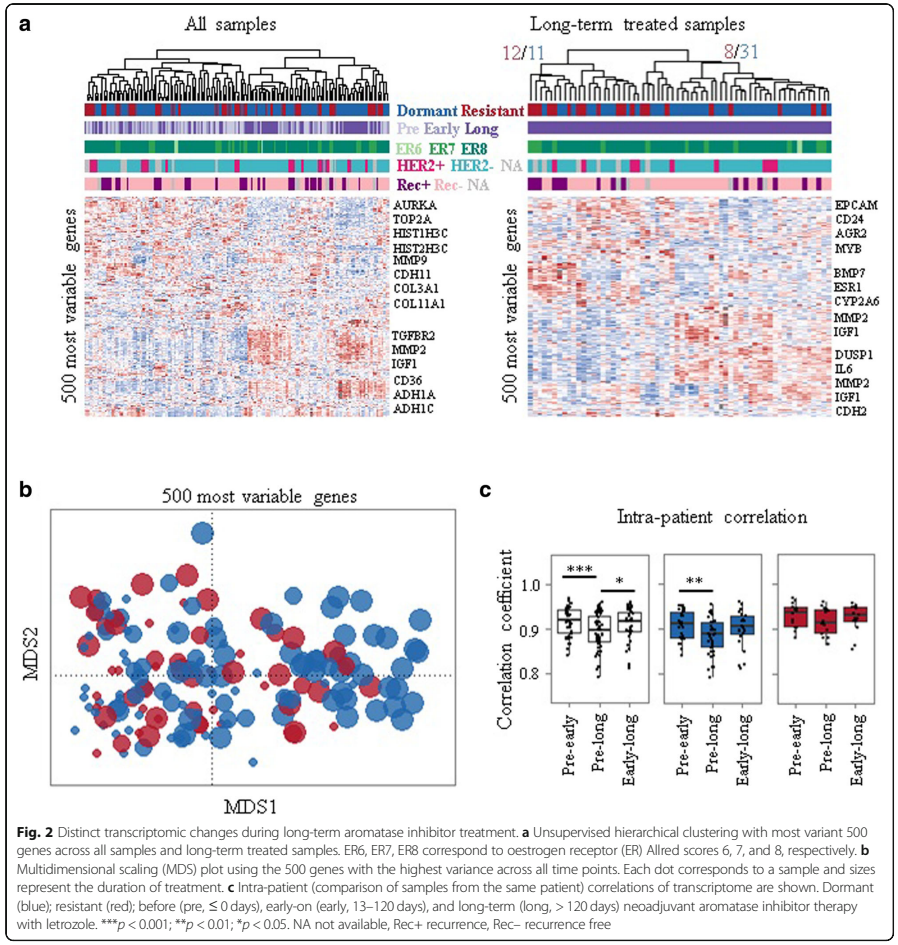

Fig. 2 Distinct transcriptomic changes during long-term aromatase inhibitor treatment. **a** Unsupervised hierarchical clustering with most variant 500 genes across all samples and long-term treated samples. ER6, ER7, ER8 correspond to oestrogen receptor (ER) Allred scores 6, 7, and 8, respectively. **b** Multidimensional scaling (MDS) plot using the 500 genes with the highest variance across all time points. Each dot corresponds to a sample and sizes represent the duration of treatment. **c** Intra-patient (comparison of samples from the same patient) correlations of transcriptome are shown. Dormant (blue); resistant (red); before (pre, ≤ 0 days), early-on (early, 13–120 days), and long-term (long, > 120 days) neoadjuvant aromatase inhibitor therapy with letrozole. ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$. NA not available, Rec+ recurrence, Rec− recurrence free

differentially expressed between long-term treated dormant and resistant tumours (data not shown).

Changes in gene expression/pathways following long-term letrozole treatment

To consider whether the gene expression changes due to treatment in the dormant and acquired resistant tumours were the same or distinct we initially considered them separately. Pairwise rank product analysis (pre- versus long-term treatment, false discovery rate (FDR) < 0.01) identified 2319 genes significantly differentially expressed (1063 downregulated and 1256 upregulated) between long-term treated and pre-treatment dormant tumours (Additional file 2: Table S2). These genes were significantly enriched ($p < 0.01$) for a total of 62 and 26 pathways, respectively (Additional file 2: Table S3), including reductions in cell cycle, senescence, DNA methylation, and an increase in extracellular matrix (ECM) organization. These findings are consistent with previous studies of patient-matched sequential samples treated with oestrogen deprivation therapy [16–18]. Acquired resistant tumours displayed much fewer consistently differentially expressed genes (238; 63 downregulated and 175 upregulated) between long-term treated and pre-treatment samples (Additional file 2: Table S4). Genes that were upregulated in resistant tumours (pre-treatment versus long-term treatment) were enriched for several of the same pathways as dormant tumours (ECM organization, elastic fibre formation, and platelet degranulation), but downregulated genes were much more variable (Additional file 2: Table S5; Fig. 3a, b).

Having determined that dormant and acquired resistant tumours have somewhat distinct changes during treatment at the molecular level, the question remained as to whether these changes tend to occur at earlier time points or were specific to long-term treatment. For dormant tumours, differential expression begins early on, but becomes more pronounced at later time points (Fig. 3a). Downregulated genes (pre-treatment versus long-term treatment) were most evident at early-on treatment for resistant tumours, consistent with their initial response to treatment, whilst upregulated genes (pre-treatment versus long-term treatment) were most changed after long-term treatment, potentially suggesting that these genes may mediate acquired resistance (Fig. 3b). We further examined whether differentially expressed genes between pre-treatment versus long-term treatment identified in each response class (dormant and resistant) were shared (Fig. 3c, d). Both downregulated and upregulated genes identified in resistant tumours were significantly changed ($p < 0.01$) in dormant tumours (Fig. 3d). However, only upregulated genes identified in dormant tumours were significantly upregulated in resistant tumours without any change in downregulated genes (Fig. 3c), implicating a partial lack of response to treatment at the molecular level in acquired resistance patients.

A potential role of epigenetic regulation in acquired resistance

The above findings suggest that therapy-induced dynamic changes in gene expression and pathways are common features of long-term treatment, rather than being specific to dormant or resistant phenotypes. This led us to perform comparative analysis of dormant and acquired resistant tumours at the long-term time point to identify any specific differences. Unpaired rank product analysis (FDR < 0.01) revealed a total of 419 genes (170 downregulated and 249 upregulated) to be differentially expressed between long-term treated dormant and resistant tumours (Additional file 2: Table S6; Fig. 4a). These genes were significantly enriched in 27 pathways ($p < 0.05$), including several epigenetics-related pathways, including "DNA methylation", "PRC2 methylates histones and DNA", "histone acetyl transferases (HATs) acetylate histones", and "epigenetic regulation of gene expression", as well as senescence and cell cycle (Additional file 2: Table S7; Fig. 4b). Examination of the expression of these genes alone demonstrated that they could partially separate dormant from the majority of resistant tumours (Fig. 4c). Single-sample gene set enrichment analysis (ssGSEA) [33] was performed to quantitatively score the activity of differentially expressed genes in every sample. The score of differentially upregulated genes between long-term treated dormant and resistant tumours was significantly higher in acquired resistant compared with dormant tumours under early treatment ($p < 0.05$) as well as long-term treatment ($p < 0.001$) (Fig. 4d).

Our results prompted us to examine whether the changes we observed in clinical samples were similarly changed in experimental models of resistant breast cancer cells. Oestrogen receptor-positive MCF7 cells stably transfected with the aromatase gene (MCF7aro cells) and long-term oestrogen-deprived (LTED) breast cancer cells have been widely used to understand mechanisms of aromatase inhibitor resistance in vitro. Examining two publicly available gene expression datasets (GSE10879 and GSE10911) demonstrated that genes differentially expressed (upregulated) between acquired resistant and dormant tumours (a total of 249) were significantly enriched in aromatase inhibitor-resistant cells compared with sensitive/control cells (Fig. 5a). A total of 211 and 174 out of 249 genes were present in GSE10879 and GSE10911, respectively. In two out of three in-vitro studies with dynamic gene expression data from LTED MCF7 cells, an initial decrease in ssGSEA scores mimicking the dormancy/responsive state was followed by a later increase representing acquired resistance (Fig. 5b), further validating our results and emphasizing the utility of these in-vitro models. Interestingly, no significant difference was observed in tamoxifen- and fulvestrant-resistant MCF7 cells compared with drug-sensitive control cells (Fig. 5c) suggesting the specificity of the results to resistance to aromatase inhibitor therapy.

In addition, proteomic analysis of a subset of samples was performed which revealed differential expression in 656 proteins (279 downregulated, 377 upregulated) between long-term treated dormant and resistant tumours (rank product; $p < 0.05$; $n = 10$; Additional file 2: Table S8; Fig. 5d). A total of 36 features including S100P and HIST2H3A (H3.2) overlapped between proteomics and transcriptomics, validating the results with a different approach.

Furthermore, differentially expressed genes were uploaded to Enricher (ENCODE Histone modification 2015 dataset) [34] to determine histone modification enrichment. Two H3 lysine methylation modifications (H3K27me3 and H3K4me1) were enriched

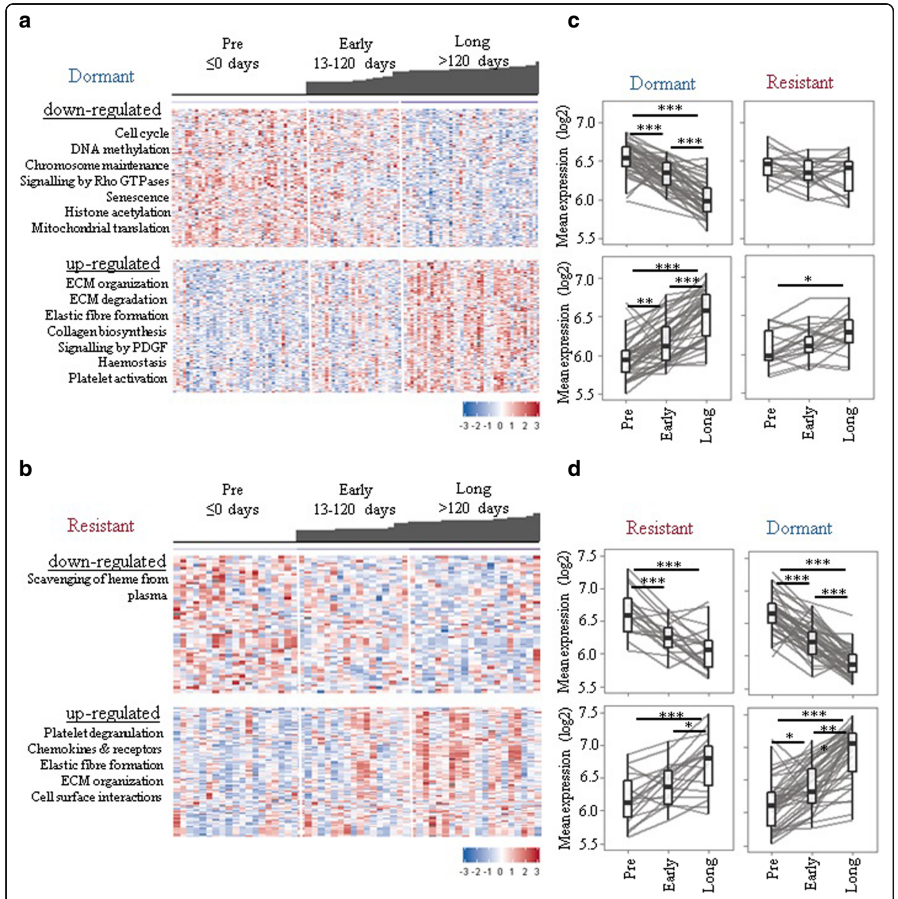

Fig. 3 Long-term oestrogen deprivation therapy is associated with cell cycle, senescence, epigenetic regulation, and extracellular matrix (ECM)-associated pathways. Differentially expressed genes between pre-treatment and long-term treated dormant (**a**) and resistant (**b**) tumours were determined. Heat-maps showing change in downregulated and upregulated gene expression in dormant (**a**) and resistant (**b**) samples. Each column represents a sample and each row a gene. Colours are log_2 mean-centred values with red indicating high values and blue indicating low expression. Bar plots on top of heat-maps represent the time on treatment (days) for each sample. **c**,**d** Graphs show dynamic changes in mean expression of differentially expressed genes in response classes. ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$

significantly (adjusted $p = 0.0003$ and $p = 0.004$, respectively) whereas no enrichment for histone acetylation was determined.

To further validate the gene expression results, immunohistochemical evaluation of FFPE sections revealed significantly lower global 5-mC and 5-hmC levels in resistant tumours compared with dormant tumours under extended treatment (Fig. 6a, b). Significantly lower 5-hmC levels in acquired resistant compared with dormant tumours were also observed at early treatment (Fig. 6b), suggesting hypomethylation may be predictive of emergence from dormancy.

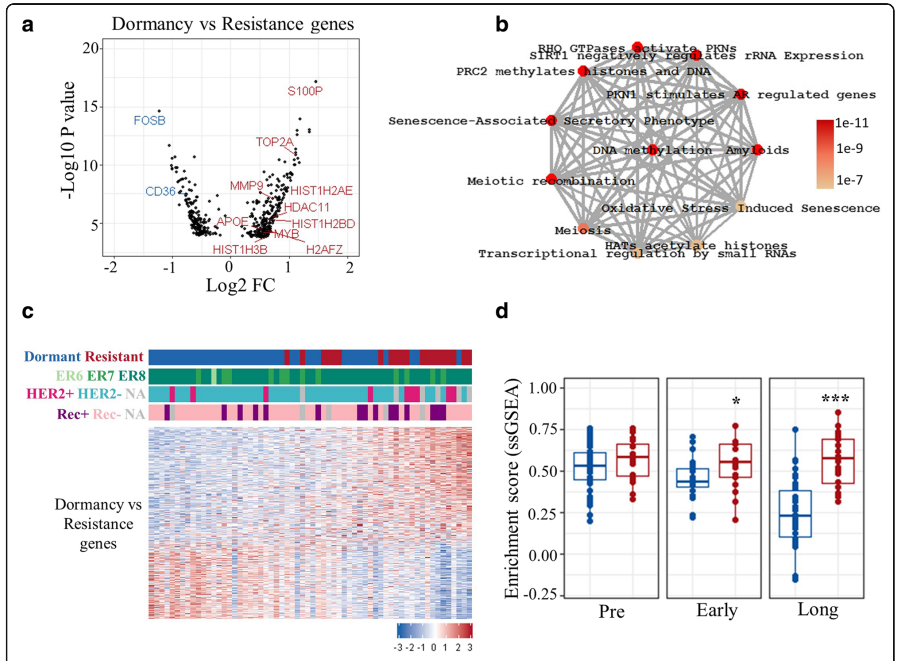

Fig. 4 Comparative analysis of dormant and resistant tumours. **a** Volcano plot showing differentially expressed genes between long-term treated dormant and resistant tumours (dormancy versus resistance genes). Some upregulated and downregulated genes in resistant tumours are highlighted in red and blue, respectively. **b** Significantly enriched pathways for dormancy versus resistance genes ($p < 0.01$; ReactomePA). Grey edges connecting the nodes indicates genes shared between the nodes/pathways, and the width of the edge is scaled by the number of common genes. Colours indicates the significance (p value) where red is a lower p value. **c** Heatmap showing partial separation of long-term treated dormant and resistant samples using dormancy versus resistance genes (a total of 419; 170 downregulated and 249 upregulated genes). Colours are log_2 mean-centred values with red indicating high and blue indicating low expression. Genes are sorted by fold-change (FC) values from most to least up/downregulated. Samples are sorted by sum expression of upregulated genes. ER6, ER7, and ER8 correspond to oestrogen receptor (ER) Allred scores 6, 7, and 8, respectively. **d** Comparison of single sample gene enrichment analysis (ssGSEA) scores of dormancy versus resistance upregulated genes between dormant and acquired resistant tumours. Dormant (blue); resistant (red); before (pre, ≤ 0 days), early-on (early, 13–120 days) and long-term (long, > 120 days) neoadjuvant aromatase inhibitor therapy with letrozole. ***$p < 0.001$; *$p < 0.05$. NA not available, Rec+ recurrence, Rec− recurrence free

Discussion

Understanding the mechanisms underlying the maintenance of and escape from dormancy have great importance considering that most cancer-related deaths are caused by metastasis rather than the primary tumour. In this study, we describe the first sequential patient-matched clinical dataset of extended oestrogen deprivation therapy in breast cancer. The results highlight the difficulty of distinguishing dormant and resistant tumours, with dynamic molecular changes of treatment being highly similar between the groups. However, comparative analysis revealed a set of genes significantly upregulated in resistant tumours compared with dormant tumours within the first months of letrozole treatment suggesting a predictive role for changes in DNA methylation.

Failure to reduce proliferation after 2 weeks of oestrogen deprivation therapy [16, 35] may well identify patients that are innately resistant; however, acquired resistance remains a greater challenge in terms of identifying biomarkers and appropriate alternative or combination therapies [36]. Many of the transcriptomic changes identified in long-term treated dormant tumours are shared by some, but not all, resistant

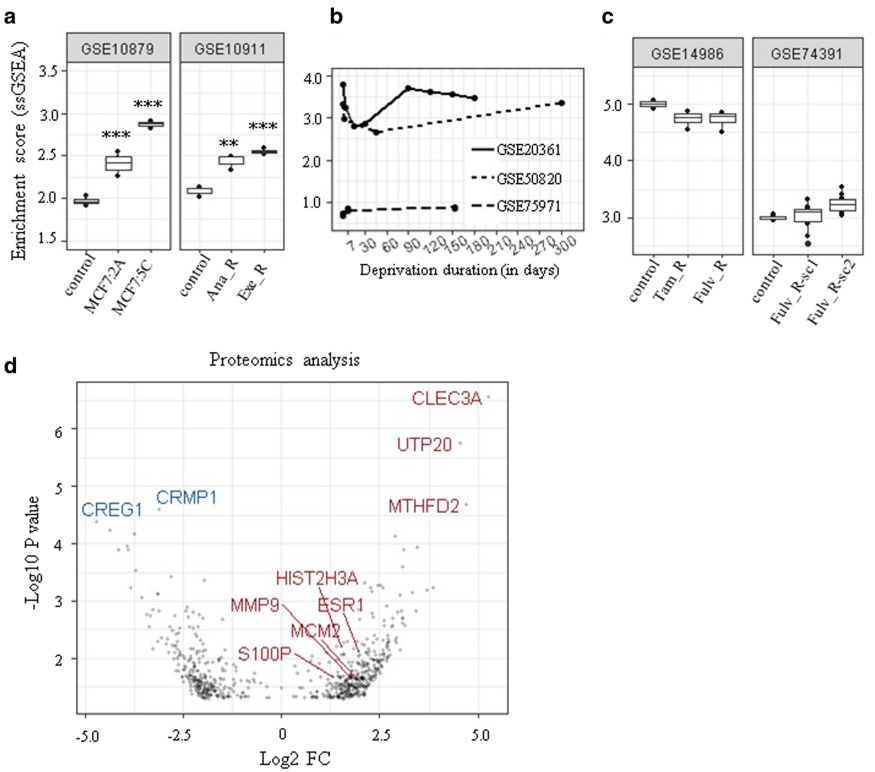

Fig. 5 Validation of results using in-vitro gene expression data from resistant cell lines and proteomics analysis. **a** Normalised enrichment scores of differently upregulated genes (a total of 249) between long-term treated dormant and resistant tumours calculated using single sample gene set enrichment analysis (ssGSEA) in aromatase inhibitor-resistant cells. Scores were significantly higher (**$p < 0.01$, ***$p < 0.001$) in two aromatase inhibitor-resistant cell lines, MCF7:2A and MCF7:5C, which were clonally derived from MCF7 breast cancer cells following long-term oestrogen deprivation (LTED) compared with control/sensitive MCF7 cells ($n = 4$). Anastrozole-resistant (Ana_R) and exemestane-resistant (Exe_R) MCF7aro cells had significantly higher scores compared with control ($n = 3$). **b** Dynamic changes in enrichment scores of LTED MCF7 cells in three different datasets. **c** Scores in tamoxifen-resistant (Tam_R) and fulvestrant-resistant (Fulv_R) and drug-sensitive (control) MCF7 cells ($n = 4$, $n = 10$). **d** Volcano plot showing differentially expressed proteins between long-term treated dormant and resistant tumours ($p < 0.05$). Some overlapping features between transcriptomics and proteomics analysis and the most upregulated and downregulated proteins are highlighted in red and blue, respectively. FC fold-change, sc subclone

tumours, providing further evidence of resistance heterogeneity [37] where dormant tumours share similar molecular changes, but there are likely to be a variety of escape mechanisms that lead to acquired resistance.

In the present study, paired differential expression analysis demonstrated that dormant tumours continue to change under long-term treatment. Some of the identified dormancy-related pathways such as cell cycle arrest and senescence have established roles in metastasis dormancy [38], further supporting the relevance of our clinical model, with the senescence-associated secretory phenotype (SASP) recently suggested to regulate breast cancer dormancy and relapse [39]. As in short-term responsive tumours [16], ECM organization and degradation were significantly upregulated in dormant tumours. ECM remodelling and its degradation by

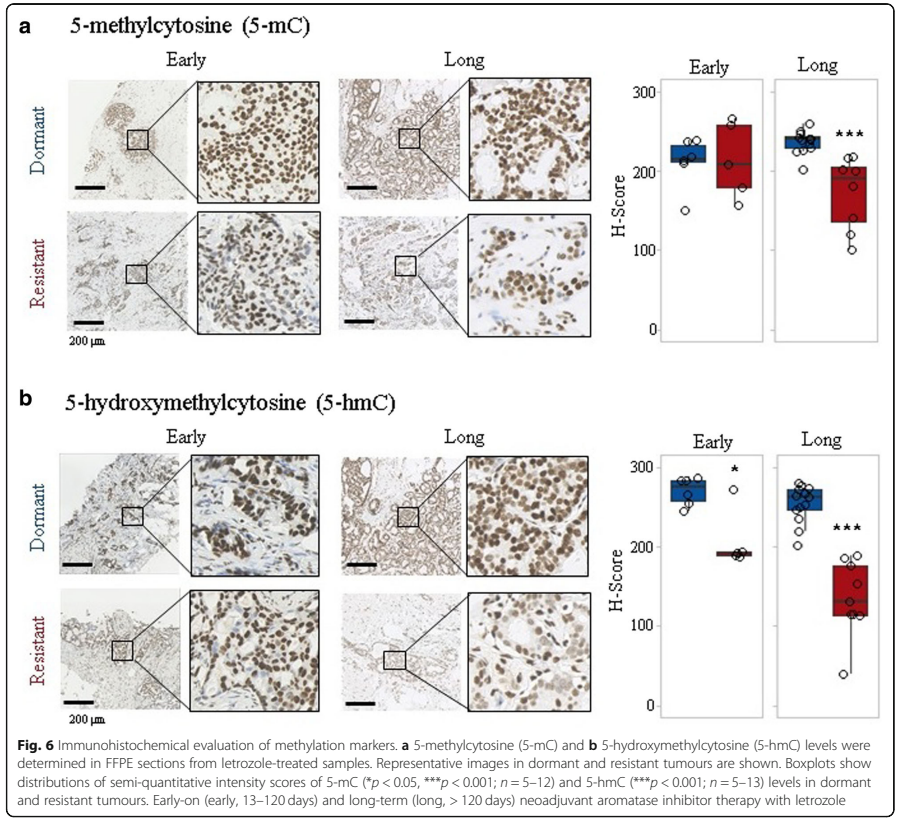

Fig. 6 Immunohistochemical evaluation of methylation markers. **a** 5-methylcytosine (5-mC) and **b** 5-hydroxymethylcytosine (5-hmC) levels were determined in FFPE sections from letrozole-treated samples. Representative images in dormant and resistant tumours are shown. Boxplots show distributions of semi-quantitative intensity scores of 5-mC (*$p < 0.05$, ***$p < 0.001$; $n = 5$–12) and 5-hmC (***$p < 0.001$; $n = 5$–13) levels in dormant and resistant tumours. Early-on (early, 13–120 days) and long-term (long, > 120 days) neoadjuvant aromatase inhibitor therapy with letrozole

matrix metalloproteases (MMP) have previously been suggested to regulate the switch between dormancy and metastatic growth [40]. Despite histological confirmation that each tumour sample contained at least 60% tumour, we acknowledge that the results presented are of intact whole tissue and potentially limited by minor variations in tumour cellularity or the proportion of stoma which could affect gene expression.

The most transcriptionally upregulated gene in resistant tumours *S100P*, previously shown to be an inducer of breast cancer metastasis correlated with decreased survival [41]. S100P, a small calcium-binding protein mediating Ca^{2+}-dependent signalling pathways, has distinct functions in normal tissue and cancer, including human embryonic development and breast cancer initiation [42]. Recently, *S100P* hypomethylation in blood was demonstrated to be inversely correlated with tissue *S100P* expression and significantly associated with breast cancer, implicating *S100P* as a potential diagnostic marker [43]. High plasma *S100P* levels have also been correlated with poor prognosis in metastatic breast cancer patients, with levels decreasing following treatment, suggesting a role of S100P in dynamic monitoring of response [44]. In the present study, S100P gene expression and protein levels were significantly higher in resistant tumours after long-term treatment, as well as being differentially expressed before treatment, supporting its potential role as a therapeutic target [45] and a predictive marker.

Comparative analysis of dormant and resistant samples after extended treatment revealed enrichment for a set of genes with a role in DNA methylation and histone

acetylation/deacetylation. Epigenetic alterations are recognized to occur in breast cancer. DNA methyltransferase (DNMT) and histone deacetylase (HDAC) inhibitors have been shown to exert encouraging effects on the disease [46]. Recently, the potential role of epigenetic changes in regulating dormancy and reactivation state has been suggested to explain the reversible (on/off) nature of dormancy [47].

Breast cancer "CpG island methylator phenotype" (CIMP), as revealed by genome-wide methylation analysis of metastatic breast cancers where a large number of genes are hypermethylated, has been suggested to be informative for metastatic potential [48]. A significant correlation between pre-treatment global DNA methylation with neoadjuvant chemotherapy response in rectal cancer has been reported [49]. Although DNA hypomethylation was the first epigenetic alteration identified in cancer, its molecular process and effects are not yet well understood [50]. In addition, 5-hmC levels were shown to correlate with differentiation status, with higher levels when more differentiated [29]. In addition, alterations in DNA methylation in LTED MCF7 cells have been previously reported [51]. Our results provide evidence for the loss of a global DNA methylation process in resistant tumours and strengthen the case to use these models for further study. The global decrease in 5-mC may account for the observed reduction in 5-hmC levels since 5-mC is converted to 5-hmC. On the other hand, at the early-on time point, 5-hmC levels were significantly reduced with no significant change in 5-mC levels, suggesting an independent role of the 5-hmC mark. Hypomethylated cancer cells have been suggested to be selected to form tumours with increased malignancy [50]. We suggest that hypomethylation in resistant tumours may reflect a de-differentiation process inducing stem cell-like cell formation. Determining the time point at which that hypomethylation starts, which would allow intervention before it starts to prevent resistance to therapy, needs further investigation.

The main genes significantly enriched for epigenetics-associated pathways in the present study are core histone (H3, H4, H2B) genes. Well-known epigenetics-associated genes such as DNMT were not differentially expressed in the present study. Therefore, it might be suggested that observed changes in histone gene levels may simply reflect the high proliferation rate in resistant tumours since transcription of these histone genes are replication-dependent and their mRNA levels increase during DNA replication [52]. However, deregulation of histone H2A and H2B was associated with anthracycline resistance in breast cancer cells and reversed by HDAC small molecule inhibitors [53]. Furthermore, upregulation of replication-dependent core histone proteins has been suggested to be a selective indicator of ER-mediated MCF7 cell proliferation regardless of the proliferation rate [54]. Also, observed global loss of DNA methylation in resistant tumours suggests dynamic regulation of gene transcription under letrozole therapy. Therefore, histone upregulation and alterations in epigenetic pathways observed in our study may play a role in resistance to endocrine deprivation therapy, rather than simply mirroring the degree of proliferation.

Our results indicate alterations both in DNA methylation and histone modifications. HDAC inhibitors, which have been shown to regulate DNA methylation [55], may be successful clinically as second-line drugs alone or in combination following oestrogen deprivation therapy failure as there is growing evidence for their tumour selective action [56, 57]. A time-dependent role for HDACs in leukaemia has been shown [58] and may also be critical in determining when to start HDAC inhibition therapy to successfully treat tumours resistant to oestrogen deprivation therapy. Whether or not the epigenetic alterations are triggers of re-awakening and if the timely use of epigenetic drugs can prevent acquired resistance warrants further investigation.

Conclusions

We have performed the first study of sequential tumour samples from breast cancer patients receiving extended neoadjuvant oestrogen deprivation therapy as a clinical model of dormancy and acquired resistance. Our analysis suggests that molecular differences between dormant and resistant tumours are initially subtle, becoming more obvious only after extended treatment. This study emphasizes that alterations in DNA methylation in the first months of treatment may predict which patients will eventually develop acquired resistance.

Additional files

Additional file 1: Figure S1. Consort diagram showing the cohort and sample sizes. Patient and sample sizes in each group are shown with inclusion and exclusion criteria. (JPG 80 kb)

Additional file 2: Table S1. Table indicating with which public dataset each sample is associated. **Table S2.** Differentially expressed genes between long-term treated and pre-treatment dormant tumours. **Table S3.** Enriched pathways for differentially expressed genes between long-term treated and pre-treatment dormant tumours. **Table S4.** Differentially expressed genes between long-term treated and pre-treatment acquired resistant tumours. **Table S5.** Enriched pathways for differentially expressed genes between long-term treated and pre-treatment acquired resistant tumours. **Table S6.** Differentially expressed genes between long-term treated dormant and resistant tumours. **Table S7.** Enriched pathways for differentially expressed genes between long-term treated dormant and resistant tumours. **Table S8.** Differentially expressed proteins between long-term treated dormant and resistant tumours. (XLSX 286 kb)

Abbreviations

5-hmC: 5-Hydroxymethylcytosine; 5-mC: 5-Methylcytosine; CIMP: CpG island methylator phenotype; DAB: 3,3'-Diaminobenzidine; DNMT: DNA methyltransferase; ECM: Extracellular matrix; ER: Oestrogen receptor; FASP: Filter Aided Sample Preparation; FDR: False discovery rate; FFPE: Formalin-fixed paraffin-embedded; HAT: Histone acetyl transferase; HDAC: Histone deacetylase; HRP: Horseradish peroxidase; LFQ: Label-free Quantitation; LTED: Long-term oestrogen deprived; MDS: Multidimensional scaling; MMP: Matrix metalloproteases; SASP: Senescence-associated secretory phenotype; ssGSEA: Single-sample gene set enrichment analysis

Acknowledgements

We are grateful for the help and support of the Genetics Team at the Edinburgh Wellcome Trust Clinical Research Facility under the direction of Lee Murphy.

Funding

The work was supported Marie Skłodowska-Curie Individual Fellowship (H2020-MSCA-IF, 658170) to CS. AHS and JMD are very grateful for funding provided by Breast Cancer Now. The work was partly supported by Welcome Trust Institutional Fund (ISSF3) to CS and AHS.

Availability of data and materials

The microarray dataset generated during the current study is available in NCBI GEO (www.ncbi.nlm.nih.gov/geo) under accession GSE111563. The analysis also includes previously published microarray data under accession numbers GSE59515 and GSE55374. The proteomics dataset generated during the current study is available in PRIDE (www.ebi.ac.uk/pride/archive) with the identifier PXD009328. Publicly available resistant cell line gene expression datasets GSE10879, GSE10911 [59], GSE20361 [60], GSE50820 [61], GSE75971 [62], GSE14986 [63], GSE74391 [64] were also analysed.

Authors' contributions

AKT, JMD, and AHS conceived the study. CS and AKT generated the transcriptome dataset. JW conducted proteomics and supported proteomics data analysis. DAP, AL, and AHS provided help with the data analysis. AF and LR provided technical support with tissue collection and processing. LR, JST, and JMD co-ordinated the collection and assessment of clinical samples. CS analysed and interpreted the data and drafted the manuscript. AHS supervised the project and helped to write and edit the manuscript. All authors read and approved the final manuscript.

Authors' information

Not applicable.

Ethics approval and consent to participate

All patients provided informed consent and sample collection was approved by the local research ethics committee (Lothian Local Research Ethics Committee 03, REC Reference number 07/S1103/26, approval date 13 August 2007).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]Applied Bioinformatics of Cancer, University of Edinburgh Cancer Research UK Centre, MRC Institute of Genetics and Molecular Medicine, Edinburgh, UK. [2]Department of Pharmacology, Faculty of Pharmacy, Ege University, 35040 Izmir, Turkey. [3]Edinburgh Breast Unit, Western General Hospital, Edinburgh, UK. [4]Mass Spectrometry Unit, MRC Institute of Genetics and Molecular Medicine, Edinburgh, UK.

Received: 10 July 2018 Accepted: 19 December 2018
Published online: 07 January 2019

References

1. EBCTCG. Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. Lancet. 2015;386(10001):1341–52.
2. Demicheli R, Ardoino I, Boracchi P, Coradini D, Agresti R, Ferraris C, et al. Recurrence and mortality according to estrogen receptor status for breast cancer patients undergoing conservative surgery. Ipsilateral breast tumour recurrence dynamics provides clues for tumour biology within the residual breast. BMC Cancer. 2010;10:656.
3. Pan HC, Gray R, Braybrooke J, Davies C, Taylor C, McGale P, et al. 20-Year risks of breast-cancer recurrence after stopping endocrine therapy at 5 years. N Engl J Med. 2017;377(19):1836–46.
4. Weigelt B, Glas AM, Wessels LF, Witteveen AT, Peterse JL, van't Veer LJ. Gene expression profiles of primary breast tumors maintained in distant metastases. Proc Natl Acad Sci U S A. 2003;100(26):15901–5.
5. Tang MH, Dahlgren M, Brueffer C, Tjitrowirjo T, Winter C, Chen Y, et al. Remarkable similarities of chromosomal rearrangements between primary human breast cancers and matched distant metastases as revealed by whole-genome sequencing. Oncotarget. 2015;6(35):37169–84.
6. Kroigard AB, Larsen MJ, Thomassen M, Kruse TA. Molecular concordance between primary breast cancer and matched metastases. Breast J. 2016;22(4):420–30.
7. Demicheli R, Terenziani M, Bonadonna G. Estimate of tumor growth time for breast cancer local recurrences: rapid growth after wake-up? Breast Cancer Res Treat. 1998;51(2):133–7.
8. Uhr JW, Pantel K. Controversies in clinical cancer dormancy. Proc Natl Acad Sci U S A. 2011;108(30):12396–400.
9. Sosa MS, Bragado P, Aguirre-Ghiso JA. Mechanisms of disseminated cancer cell dormancy: an awakening field. Nat Rev Cancer. 2014;14(9):611–22.
10. Dittmer J. Mechanisms governing metastatic dormancy in breast cancer. Semin Cancer Biol. 2017;44:72–82.
11. Selli C, Dixon JM, Sims AH. Accurate prediction of response to endocrine therapy in breast cancer patients: current and future biomarkers. Breast Cancer Res. 2016;18(1):118.
12. Clarke R, Tyson JJ, Dixon JM. Endocrine resistance in breast cancer—an overview and update. Mol Cell Endocrinol. 2015;418(Pt 3):220–34.
13. Ma CX, Reinert T, Chmielewska I, Ellis MJ. Mechanisms of aromatase inhibitor resistance. Nat Rev Cancer. 2015;15(5):261–75.
14. Sims AH, Bartlett JMS. Approaches towards expression profiling the response to treatment. Breast Cancer Res. 2008;10(6):115.
15. Vendrell JA, Robertson KE, Ravel P, Bray SE, Bajard A, Purdie CA, et al. A candidate molecular signature associated with tamoxifen failure in primary breast cancer. Breast Cancer Res. 2008;10(5):R88.
16. Turnbull AK, Arthur LM, Renshaw L, Larionov AA, Kay C, Dunbier AK, et al. Accurate prediction and validation of response to endocrine therapy in breast cancer. J Clin Oncol. 2015;33(20):2270–8.
17. Dunbier AK, Ghazoui Z, Anderson H, Salter J, Nerurkar A, Osin P, et al. Molecular profiling of aromatase inhibitor-treated postmenopausal breast tumors identifies immune-related correlates of resistance. Clin Cancer Res. 2013;19(10):2775–86.
18. Patani N, Dunbier AK, Anderson H, Ghazoui Z, Ribas R, Anderson E, et al. Differences in the transcriptional response to fulvestrant and estrogen deprivation in ER-positive breast cancer. Clin Cancer Res. 2014;20(15):3962–73.
19. Arthur LM, Turnbull AK, Webber VL, Larionov AA, Renshaw L, Kay C, et al. Molecular changes in lobular breast cancers in response to endocrine therapy. Cancer Res. 2014;74(19):5371–6.
20. Turnbull AK, Kitchen RR, Larionov AA, Renshaw L, Dixon JM, Sims AH. Direct integration of intensity-level data from Affymetrix and Illumina microarrays improves statistical power for robust reanalysis. BMC Med Genet. 2012;5:35.
21. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. Bioinformatics. 2008;24(13):1547–8.
22. Yu G, He QY. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. Mol BioSyst. 2016;12(2):477–9.
23. Hong FX, Breitling R, McEntee CW, Wittner BS, Nemhauser JL, Chory J. RankProd: a bioconductor package for detecting differentially expressed genes in meta-analysis. Bioinformatics. 2006;22(22):2825–7.

7 | References

1. Pearce DA, Arthur LM, Turnbull AK, Renshaw L, Sabine VS, Thomas JS, et al. Tumour sampling method can significantly influence gene expression profiles derived from neoadjuvant window studies. Scientific Reports [Internet]. 2016 Jul 7 [cited 2018 Apr 26];6:29434. Available from: https://www.nature.com/articles/srep29434

2. Selli C, Turnbull AK, Pearce DA, Li A, Fernando A, Wills J, et al. Molecular changes during extended neoadjuvant letrozole treatment of breast cancer: Distinguishing acquired resistance from dormant tumours. Breast Cancer Research [Internet]. 2019 Jan;21(1). Available from: https://doi.org/10.1186%2Fs13058-018-1089-5

3. Pearce DA, Nirmal AJ, Freeman TC, Sims AH. Continuous biomarker assessment by exhaustive survival analysis. 2017 Oct; Available from: https://doi.org/10.1101%2F208660

4. Harold E. Varmus - Banquet Speech [Internet]. [cited 2018 Mar 13]. Available from: https://www.nobelprize.org/nobel_prizes/medicine/laureates/1989/varmus-speech.html

5. Domazet-Lošo T, Klimovich A, Anokhin B, Anton-Erxleben F, Hamm MJ, Lange C, et al. Naturally occurring tumours in the basal metazoan *Hydra*. Nature Communications [Internet]. 2014 Jun 24 [cited 2018 Mar 16];5:4222. Available from: https://www.nature.com/articles/ncomms5222

6. Cairns J. Mutation selection and the natural history of cancer. Nature [Internet]. 1975 May [cited 2018 Mar 16];255(5505):197–200. Available from: https://www.nature.com/articles/255197a0

7. Hanahan D, Weinberg RA. The Hallmarks of Cancer. Cell [Internet]. 2000 Jan 7 [cited 2018 Mar 16];100(1):57–70. Available from: http://www.sciencedirect.com/science/article/pii/S0092867400816839

8. Cheung-Ong K, Giaever G, Nislow C. DNA-Damaging Agents in Cancer Chemotherapy: Serendipity and Chemical Biology. Chemistry & Biology [Internet]. 2013 May [cited 2018 Mar 16];20(5):648–59. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1074552113001312

9. Raguz S, Yagüe E. Resistance to chemotherapy: New treatments and novel insights into an old problem. British Journal of Cancer [Internet]. 2008 Aug [cited 2018 Mar 16];99(3):387–91. Available from: https://www.nature.com/articles/6604510

10. Ng CKY, Schultheis AM, Bidard F-C, Weigelt B, Reis-Filho JS. Breast cancer genomics from microarrays to massively parallel sequencing: Paradigms and new insights. Journal of the National Cancer Institute [Internet]. 2015 May;107(5):djv015. Available from: http://jnci.oxfordjournals.org/content/107/5/djv015.

full

11. De Souza YG, Greenspan JS. Biobanking Past, Present and Future: Responsibilities and Benefits. AIDS [Internet]. 2013 Jan 28 [cited 2018 Mar 16];27(3):303–12. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3894636/

12. Breast cancer statistics [Internet]. Cancer Research UK. 2015 [cited 2018 Mar 16]. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer

13. Becker S. A historic and scientific review of breast cancer: The next global healthcare challenge. International Journal of Gynecology & Obstetrics [Internet]. 2015 Sep;131:S36–9. Available from: https://doi.org/10.1016%2Fj.ijgo.2015.03.015

14. Sinn H-P, Kreipe H. A brief overview of the WHO classification of breast tumors, 4th edition, focusing on issues and updates from the 3rd edition. Breast Care [Internet]. 2013;8(2):149–54. Available from: https://doi.org/10.1159%2F000350774

15. Galea MH, Blamey RW, Elston CE, Ellis IO. The Nottingham Prognostic Index in primary breast cancer. Breast Cancer Res Treat. 1992;22(3):207–19.

16. Olivotto IA, Bajdik CD, Ravdin PM, Speers CH, Coldman AJ, Norris BD, et al. Population-based validation of the prognostic model ADJUVANT! For early breast cancer. J Clin Oncol. 2005 Apr 20;23(12):2716–25.

17. Ravdin PM, Siminoff LA, Davis GJ, Mercer MB, Hewlett J, Gerson N, et al. Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. J Clin Oncol. 2001 Feb 15;19(4):980–91.

18. Chang TW, Kuo YL. A model building exercise of mortality risk for Taiwanese women with breast cancer. BMC Med Inform Decis Mak. 2010 Aug 19;10:43.

19. Cancer IA for R on, Organization WH. WHO Classification of Tumours of the Breast. International Agency for Research on Cancer; 2012. 240 pp.

20. Osborne CK, Yochmowitz MG, Knight WA, McGuire WL. The value of estrogen and progesterone receptors in the treatment of breast cancer. Cancer. 1980 Dec 15;46(12 Suppl):2884–8.

21. Hynes NE, Stern DF. The biology of erbB-2/nue/HER-2 and its role in cancer. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer [Internet]. 1994 Dec 30 [cited 2018 Mar 19];1198(2):165–84. Available from: http://www.sciencedirect.com/science/article/pii/0304419X94900124

22. Hammond MEH, Hayes DF, Dowsett M, Allred DC, Hagerty KL, Badve S, et al. American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. J Clin Oncol. 2010 Jun 1;28(16):2784–95.

23. Wolff AC, Hammond MEH, Hicks DG, Dowsett M, McShane LM, Allison KH, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. J Clin Oncol. 2013 Nov 1;31(31):3997–4013.

24. Beatson GT. On the Treatment of Inoperable Cases of Carcinoma of the Mamma: Suggestions for a

New Method of Treatment, with Illustrative Cases. Transactions Medico-Chirurgical Society of Edinburgh [Internet]. 1896;15:153–79. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29584099 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5518378

25. Yersal O, Barutca S. Biological subtypes of breast cancer: Prognostic and therapeutic implications. World J Clin Oncol. 2014 Aug 10;5(3):412–24.

26. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: An overview of the randomised trials. Lancet. 2005May14–20AD;365(9472):1687–717.

27. Early Breast Cancer Trialists' Collaborative Group (EBCTCG), Davies C, Godwin J, Gray R, Clarke M, Cutter D, et al. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: Patient-level meta-analysis of randomised trials. Lancet. 2011 Aug 27;378(9793):771–84.

28. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012 Oct 4;490(7418):61–70.

29. Gruvberger S, Ringnér M, Chen Y, Panavally S, Saal LH, Borg A null, et al. Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns. Cancer Res. 2001 Aug 15;61(16):5979–84.

30. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. Serial analysis of gene expression. Science [Internet]. 1995 Oct;270(5235):484–7. Available from: https://doi.org/10.1126%2Fscience.270.5235.484

31. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science [Internet]. 1995 Oct;270(5235):467–70. Available from: https://doi.org/10.1126%2Fscience.270.5235.467

32. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature [Internet]. 2000 Aug;406(6797):747–52. Available from: http://dx.doi.org/10.1038/35021093

33. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science (New York, NY) [Internet]. 1999 Oct;286(5439):531–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10521349

34. Hu Z, Fan C, Oh DS, Marron J, He X, Qaqish BF, et al. The molecular portraits of breast tumors are conserved across microarray platforms. BMC Genomics. 2006;12.

35. Parker JS, Mullins M, Cheang MCU, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. Journal of clinical oncology : official journal of the American Society of Clinical Oncology [Internet]. 2009 Mar;27(8):1160–7. Available from: http://jco.ascopubs.org/content/27/8/1160

36. Hennigs A, Riedel F, Gondos A, Sinn P, Schirmacher P, Marmé F, et al. Prognosis of breast cancer molecular subtypes in routine clinical care: A large prospective cohort study. BMC Cancer [Internet]. 2016 Sep 15 [cited 2018 Mar 19];16. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5024419/

37. Kennecke H, Yerushalmi R, Woods R, Cheang MCU, Voduc D, Speers CH, et al. Metastatic behavior of breast cancer subtypes. Journal of Clinical Oncology [Internet]. 2010 Jul;28(20):3271–7. Available from:

https://doi.org/10.1200%2Fjco.2009.25.9820

38. Mackay A, Weigelt B, Grigoriadis A, Kreike B, Natrajan R, A'Hern R, et al. Microarray-based class discovery for molecular classification of breast cancer: Analysis of interobserver agreement. Journal of the National Cancer Institute [Internet]. 2011 Apr;103(8):662–73. Available from: http://jnci.oxfordjournals.org/content/103/8/662.abstract

39. Llombart-Cussac A, Cortés J, Paré L, Galván P, Bermejo B, Martínez N, et al. HER2-enriched subtype as a predictor of pathological complete response following trastuzumab and lapatinib without chemotherapy in early-stage HER2-positive breast cancer (PAMELA): An open-label, single-group, multicentre, phase 2 trial. The Lancet Oncology [Internet]. 2017 Apr;18(4):545–54. Available from: https://doi.org/10.1016%2Fs1470-2045%2817%2930021-9

40. Gianni L, Eiermann W, Semiglazov V, Lluch A, Tjulandin S, Zambetti M, et al. Neoadjuvant and adjuvant trastuzumab in patients with HER2-positive locally advanced breast cancer (NOAH): Follow-up of a randomised controlled superiority trial with a parallel HER2-negative cohort. The Lancet Oncology [Internet]. 2014 May;15(6):640–7. Available from: https://doi.org/10.1016%2Fs1470-2045%2814%2970080-4

41. Smid M, Wang Y, Zhang Y, Sieuwerts AM, Yu J, Klijn JG, et al. Subtypes of breast cancer show preferential site of relapse. Cancer Research [Internet]. 2008 May;68(9):3108–14. Available from: https://doi.org/10.1158%2F0008-5472.can-07-5644

42. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biology [Internet]. 2007;8(5):R76. Available from: https://doi.org/10.1186%2Fgb-2007-8-5-r76

43. Curtis C, Shah SP, Chin S-F, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature [Internet]. 2012 Jun;486(7403):346–52. Available from: http://dx.doi.org/10.1038/nature10983

44. Ali HR, Rueda OM, Chin S-F, Curtis C, Dunning MJ, Aparicio SA, et al. Genome-driven integrated classification of breast cancer validated in over 7,500 samples. Genome Biology [Internet]. 2014 Aug;15(8):431. Available from: http://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0431-1

45. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci USA. 2001 Sep 11;98(19):10869–74.

46. Cheang MCU, Chia SK, Voduc D, Gao D, Leung S, Snider J, et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. J Natl Cancer Inst. 2009 May 20;101(10):736–50.

47. Sotiriou C, Neo S-Y, McShane LM, Korn EL, Long PM, Jazaeri A, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proc Natl Acad Sci USA. 2003 Sep 2;100(18):10393–8.

48. Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M, Anderson K, et al. Breast cancer molecular subtypes respond differently to preoperative chemotherapy. Clin Cancer Res. 2005 Aug 15;11(16):5678–85.

49. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AAM, Mao M, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature [Internet]. 2002 Jan;415(6871):530–6. Available from: http:

//dx.doi.org/10.1038/415530a

50. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AAM, Voskuil DW, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med. 2002 Dec 19;347(25):1999–2009.

51. Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. J Natl Cancer Inst. 2006 Sep 6;98(17):1183–92.

52. Drukker CA, Bueno-de-Mesquita JM, Retèl VP, van Harten WH, van Tinteren H, Wesseling J, et al. A prospective evaluation of a breast cancer prognosis signature in the observational RASTER study. Int J Cancer. 2013 Aug 15;133(4):929–36.

53. Cardoso F, van't Veer LJ, Bogaerts J, Slaets L, Viale G, Delaloge S, et al. 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. N Engl J Med. 2016 Aug 25;375(8):717–29.

54. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. The New England journal of medicine. 2004 Dec;351(27):2817–26.

55. Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. J Clin Oncol. 2006 Aug 10;24(23):3726–34.

56. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, et al. Prospective Validation of a 21-Gene Expression Assay in Breast Cancer. N Engl J Med. 2015 Nov 19;373(21):2005–14.

57. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, et al. Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. New England Journal of Medicine [Internet]. 2018 Jul;379(2):111–21. Available from: https://doi.org/10.1056%2Fnejmoa1804710

58. Goldstein LJ, Gray R, Badve S, Childs BH, Yoshizawa C, Rowley S, et al. Prognostic utility of the 21-gene assay in hormone receptor-positive operable breast cancer compared with classical clinicopathologic features. J Clin Oncol. 2008 Sep 1;26(25):4063–71.

59. Nielsen TO, Parker JS, Leung S, Voduc D, Ebbert M, Vickery T, et al. A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. Clin Cancer Res. 2010 Nov 1;16(21):5222–32.

60. Gnant M, Filipits M, Greil R, Stoeger H, Rudas M, Bago-Horvath Z, et al. Predicting distant recurrence in receptor-positive breast cancer patients with limited clinicopathological risk: Using the PAM50 Risk of Recurrence score in 1478 postmenopausal patients of the ABCSG-8 trial treated with adjuvant endocrine therapy alone. Ann Oncol. 2014 Feb;25(2):339–45.

61. Filipits M, Rudas M, Jakesz R, Dubsky P, Fitzal F, Singer CF, et al. A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. Clinical cancer research : an official journal of the American Association for Cancer Research. 2011 Sep;17(18):6012–20.

62. Kronenwett R, Bohmann K, Prinzler J, Sinn BV, Haufe F, Roth C, et al. Decentral gene expression analysis: Analytical validation of the Endopredict genomic multianalyte breast cancer prognosis test. BMC Cancer.

2012 Oct 5;12:456.

63. Dubsky P, Filipits M, Jakesz R, Rudas M, Singer CF, Greil R, et al. EndoPredict improves the prognostic classification derived from common clinical guidelines in ER-positive, HER2-negative early breast cancer. Ann Oncol. 2013 Mar;24(3):640–7.

64. Martin M, Brase JC, Calvo L, Krappmann K, Ruiz-Borrego M, Fisch K, et al. Clinical validation of the EndoPredict test in node-positive, chemotherapy-treated ER+/HER2- breast cancer patients: Results from the GEICAM 9906 trial. Breast Cancer Res. 2014 Apr 12;16(2):R38.

65. Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, et al. Gene expression profiling in breast cancer: Understanding the molecular basis of histologic grade to improve prognosis. Journal of the National Cancer Institute [Internet]. 2006 Feb;98(4):262–72. Available from: http://jnci.oxfordjournals.org/content/98/4/262.abstract

66. Fan C, Oh DS, Wessels L, Weigelt B, Nuyten DSA, Nobel AB, et al. Concordance among gene-expression-based predictors for breast cancer. N Engl J Med. 2006 Aug 10;355(6):560–9.

67. The MicroArray quality control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models. Nature Biotechnology [Internet]. 2010 Jul;28(8):827–38. Available from: https://doi.org/10.1038%2Fnbt.1665

68. The MicroArray quality control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. Nature Biotechnology [Internet]. 2006 Sep;24(9):1151–61. Available from: https://doi.org/10.1038%2Fnbt1239

69. Tofigh A, Suderman M, Paquet ER, Livingstone J, Bertos N, Saleh SM, et al. The prognostic ease and difficulty of invasive breast carcinoma. Cell reports [Internet]. 2014 Oct;9(1):129–42. Available from: http://www.sciencedirect.com/science/article/pii/S2211124714007657

70. Gruber CJ, Gruber DM, Gruber IM, Wieser F, Huber JC. Anatomy of the estrogen response element. Trends in Endocrinology & Metabolism [Internet]. 2004 Mar;15(2):73–8. Available from: https://doi.org/10.1016%2Fj.tem.2004.01.008

71. Iwamoto T, Pusztai L. Predicting prognosis of breast cancer with gene signatures: Are we lost in a sea of data? Genome Medicine [Internet]. 2010;2(11):81. Available from: https://doi.org/10.1186%2Fgm202

72. Venet D, Dumont JE, Detours V. Most random gene expression signatures are significantly associated with breast cancer outcome. PLoS computational biology [Internet]. 2011 Oct;7(10):e1002240. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22028643 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3197658

73. Shimoni Y. Association between expression of random gene sets and survival is evident in multiple cancer types and may be explained by sub-classification. Zhang B, editor. PLOS Computational Biology [Internet]. 2018 Feb;14(2):e1006026. Available from: https://doi.org/10.1371%2Fjournal.pcbi.1006026

74. Kitchen RR, Sabine VS, Simen AA, Dixon JM, Bartlett JM, Sims AH. Relative impact of key sources of systematic noise in affymetrix and illumina gene-expression microarray experiments. BMC Genomics [Internet]. 2011 Dec;12(1). Available from: https://doi.org/10.1186%2F1471-2164-12-589

75. Zhao S, Fung-Leung W-P, Bittner A, Ngo K, Liu X. Comparison of RNA-seq and microarray in transcriptome profiling of activated t cells. Zhang S-D, editor. PLoS ONE [Internet]. 2014 Jan;9(1):e78644. Available from:

https://doi.org/10.1371%2Fjournal.pone.0078644

76. Ye Q-H, Qin L-X, Forgues M, He P, Kim JW, Peng AC, et al. Predicting hepatitis B virus–positive metastatic hepatocellular carcinomas using gene expression profiling and supervised machine learning. Nature Medicine [Internet]. 2003 Apr [cited 2018 Apr 12];9(4):416–23. Available from: https://www.nature.com/articles/nm843

77. Roessler S, Lin G, Forgues M, Budhu A, Hoover S, Simpson RM, et al. Integrative Genomic and Transcriptomic Characterization of Matched Primary and Metastatic Liver and Colorectal Carcinoma. Int J Biol Sci [Internet]. 2015 Jan 1 [cited 2018 Apr 12];11(1):88–98. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4278258/

78. Stuart-Harris R, Caldas C, Pinder SE, Pharoah P. Proliferation markers and survival in early breast cancer: A systematic review and meta-analysis of 85 studies in 32,825 patients. The Breast [Internet]. 2008 Aug 1 [cited 2018 Apr 12];17(4):323–34. Available from: http://www.sciencedirect.com/science/article/pii/S0960977608000593

79. Wirapati P, Sotiriou C, Kunkel S, Farmer P, Pradervand S, Haibe-Kains B, et al. Meta-analysis of gene-expression profiles in breast cancer: Toward a unified understanding of breast cancer subtyping and prognosis signatures. Breast Cancer Research [Internet]. 2008;10(4):R65. Available from: http://breast-cancer-research.com/content/10/4/R65

80. Planche A, Bacac M, Provero P, Fusco C, Delorenzi M, Stehle J-C, et al. Identification of Prognostic Molecular Features in the Reactive Stroma of Human Breast and Prostate Cancer. PLOS ONE [Internet]. 2011 May 18 [cited 2018 Apr 12];6(5):e18640. Available from: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0018640

81. Bacac M, Provero P, Mayran N, Stehle J-C, Fusco C, Stamenkovic I. A Mouse Stromal Response to Tumor Invasion Predicts Prostate and Breast Cancer Patient Survival. PLOS ONE [Internet]. 2006 Dec 20 [cited 2018 Apr 12];1(1):e32. Available from: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0000032

82. McLendon R, Friedman A, Bigner D, Meir EGV, Brat DJ, Mastrogianakis GM, et al. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature [Internet]. 2008 Sep;455(7216):1061–8. Available from: https://doi.org/10.1038%2Fnature07385

83. Murugaesu N, Wilson GA, Birkbak NJ, Watkins TBK, McGranahan N, Kumar S, et al. Tracking the genomic evolution of esophageal adenocarcinoma through neoadjuvant chemotherapy. Cancer Discovery [Internet]. 2015 May;5(8):821–31. Available from: https://doi.org/10.1158%2F2159-8290.cd-15-0412

84. Findlay JM, Castro-Giner F, Makino S, Rayner E, Kartsonaki C, Cross W, et al. Differential clonal evolution in oesophageal cancers in response to neo-adjuvant chemotherapy. Nature Communications [Internet]. 2016 Apr;7:11111. Available from: https://doi.org/10.1038%2Fncomms11111

85. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer genome landscapes. Science [Internet]. 2013 Mar;339(6127):1546–58. Available from: https://doi.org/10.1126%2Fscience.1235122

86. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. Cell [Internet]. 2011 Mar;144(5):646–74. Available from: https://doi.org/10.1016%2Fj.cell.2011.02.013

87. Ding L, Ley TJ, Larson DE, Miller CA, Koboldt DC, Welch JS, et al. Clonal evolution in relapsed acute

myeloid leukaemia revealed by whole-genome sequencing. Nature [Internet]. 2012 Jan;481(7382):506–10. Available from: https://doi.org/10.1038%2Fnature10738

88. Nowell P. The clonal evolution of tumor cell populations. Science [Internet]. 1976 Oct;194(4260):23–8. Available from: https://doi.org/10.1126%2Fscience.959840

89. Nagy JD, Victor EM, Cropper JH. Why dont all whales have cancer? A novel hypothesis resolving petos paradox. Integrative and Comparative Biology [Internet]. 2007 May;47(2):317–28. Available from: https://doi.org/10.1093%2Ficb%2Ficm062

90. Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins TB, Veeriah S, et al. Tracking the evolution of nonSmall-cell lung cancer. New England Journal of Medicine [Internet]. 2017 Jun;376(22):2109–21. Available from: https://doi.org/10.1056%2Fnejmoa1616288

91. Russo M, Siravegna G, Blaszkowsky LS, Corti G, Crisafulli G, Ahronian LG, et al. Tumor heterogeneity and lesion-specific response to targeted therapy in colorectal cancer. Cancer Discovery [Internet]. 2015 Dec;6(2):147–53. Available from: https://doi.org/10.1158%2F2159-8290.cd-15-1283

92. Comprehensive molecular portraits of human breast tumours. Nature [Internet]. 2012 Oct;490(7418):61–70. Available from: http://dx.doi.org/10.1038/nature11412

93. Morris LG, Riaz N, Desrichard A, Şenbabaoğlu Y, Hakimi AA, Makarov V, et al. Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival. Oncotarget [Internet]. 2016 Jan;7(9). Available from: https://doi.org/10.18632%2Foncotarget.7067

94. Armenia J, Stephanie A. M. Wankowicz, Liu D, Gao J, Kundra R, Reznik E, et al. The long tail of oncogenic drivers in prostate cancer. Nature Genetics [Internet]. 2018 Apr;50(5):645–51. Available from: https://doi.org/10.1038%2Fs41588-018-0078-z

95. Sequist LV, Waltman BA, Dias-Santagata D, Digumarthy S, Turke AB, Fidias P, et al. Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. Science Translational Medicine [Internet]. 2011 Mar;3(75):75ra26–6. Available from: https://doi.org/10.1126%2Fscitranslmed.3002003

96. Hata A, Katakami N, Yoshioka H, Kaji R, Masago K, Fujita S, et al. Spatiotemporal t790m heterogeneity in individual patients with EGFR-mutant nonSmall-cell lung cancer after acquired resistance to EGFR-TKI. Journal of Thoracic Oncology [Internet]. 2015 Nov;10(11):1553–9. Available from: https://doi.org/10.1097%2Fjto.0000000000000647

97. Carreira S, Romanel A, Goodall J, Grist E, Ferraldeschi R, Miranda S, et al. Tumor clone dynamics in lethal prostate cancer. Science Translational Medicine [Internet]. 2014 Sep;6(254):254ra125–5. Available from: https://doi.org/10.1126%2Fscitranslmed.3009448

98. Magbanua MJM, Wolf DM, Yau C, Davis SE, Crothers J, Au A, et al. Serial expression analysis of breast tumors during neoadjuvant chemotherapy reveals changes in cell cycle and immune pathways associated with recurrence and response. Breast Cancer Research [Internet]. 2015 May;17(1):73. Available from: http://breast-cancer-research.com/content/17/1/73

99. Brandão RD, Veeck J, Van de Vijver KK, Lindsey P, de Vries B, van Elssen CHMJ, et al. A randomised controlled phase II trial of pre-operative celecoxib treatment reveals anti-tumour transcriptional response in primary breast cancer. Breast cancer research : BCR [Internet]. 2013 Jan;15(2):R29. Available from:

http://breast-cancer-research.com/content/15/2/R29

100. Sabine VS, Sims AH, Macaskill EJ, Renshaw L, Thomas JS, Dixon JM, et al. Gene expression profiling of response to mTOR inhibitor everolimus in pre-operatively treated post-menopausal women with oestrogen receptor-positive breast cancer. Breast Cancer Research and Treatment [Internet]. 2010 May;122(2):419–28. Available from: http://link.springer.com/10.1007/s10549-010-0928-6

101. Smith IE, Walsh G, Skene A, Llombart A, Mayordomo JI, Detre S, et al. A phase II placebo-controlled trial of neoadjuvant anastrozole alone or with gefitinib in early breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology [Internet]. 2007 Sep;25(25):3816–22. Available from: http://jco.ascopubs.org/content/25/25/3816

102. Turnbull AK, Arthur LM, Renshaw L, Larionov AA, Kay C, Dunbier AK, et al. Accurate Prediction and Validation of Response to Endocrine Therapy in Breast Cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology [Internet]. 2015 Jul;33(20):2270–8. Available from: http://jco.ascopubs.org/content/33/20/2270.full

103. Arthur LM, Turnbull AK, Webber VL, Larionov AA, Renshaw L, Kay C, et al. Molecular changes in lobular breast cancers in response to endocrine therapy. Cancer Research [Internet]. 2014 Aug;74(19):5371–6. Available from: https://doi.org/10.1158%2F0008-5472.can-14-0620

104. Miller WR, Larionov A, Renshaw L, Anderson TJ, Walker JR, Krause A, et al. Gene expression profiles differentiating between breast cancers clinically responsive or resistant to letrozole. Journal of Clinical Oncology [Internet]. 2009 Mar;27(9):1382–7. Available from: https://doi.org/10.1200%2Fjco.2008.16.8849

105. Miller WR, Larionov A, Miller W, Anderson T, Iqbal S, Dixon J, et al. Changes in expression of oestrogen regulated and proliferation genes with neoadjuvant treatment highlight heterogeneity of clinical resistance to the aromatase inhibitor, letrozole. Breast Cancer Research [Internet]. 2010;12(4):R52. Available from: http://breast-cancer-research.biomedcentral.com/articles/10.1186/bcr2611

106. Harvell DME, Spoelstra NS, Singh M, McManaman JL, Finlayson C, Phang T, et al. Molecular signatures of neoadjuvant endocrine therapy for breast cancer: Characteristics of response or intrinsic resistance. Breast Cancer Research and Treatment [Internet]. 2008 Mar;112(3):475–88. Available from: https://doi.org/10.1007%2Fs10549-008-9897-4

107. Patani N, Dunbier AK, Anderson H, Ghazoui Z, Ribas R, Anderson E, et al. Differences in the transcriptional response to fulvestrant and estrogen deprivation in ER-positive breast cancer. Clinical cancer research : an official journal of the American Association for Cancer Research [Internet]. 2014 Aug;20(15):3962–73. Available from: http://clincancerres.aacrjournals.org/content/20/15/3962.abstract

108. Semiglazov VF, Semiglazov VV, Dashyan GA, Ziltsova EK, Ivanov VG, Bozhok AA, et al. Phase 2 randomized trial of primary endocrine therapy versus chemotherapy in postmenopausal patients with estrogen receptor-positive breast cancer. Cancer [Internet]. 2007;110(2):244–54. Available from: https://doi.org/10.1002%2Fcncr.22789

109. Cardoso F, Costa A, Norton L, Senkus E, Aapro M, André F, et al. ESO-ESMO 2nd international consensus guidelines for advanced breast cancer (ABC2). The Breast [Internet]. 2014 Oct;23(5):489–502. Available from: https://doi.org/10.1016%2Fj.breast.2014.08.009

110. Doisneau-Sixou S. Estrogen and antiestrogen regulation of cell cycle progression in breast cancer cells.

Endocrine Related Cancer [Internet]. 2003 Jun;10(2):179–86. Available from: https://doi.org/10.1677%2Ferc.0.0100179

111. Jordan VC, Robinson SP. Species-specific pharmacology of antiestrogens: Role of metabolism. Fed Proc. 1987 Apr;46(5):1870–4.

112. Hu R, Hilakivi-Clarke L, Clarke R. Molecular mechanisms of tamoxifen-associated endometrial cancer (Review). Oncology Letters [Internet]. 2015 Apr 1 [cited 2018 Apr 6];9(4):1495–501. Available from: http://www.spandidos-publications.com/ol/9/4/1495/abstract

113. Wakeling A. Similarities and distinctions in the mode of action of different classes of antioestrogens. Endocrine Related Cancer [Internet]. 2000 Apr;7(1):17–28. Available from: https://doi.org/10.1677%2Ferc.0.0070017

114. Howell A, Robertson JF, Abram P, Lichinitser MR, Elledge R, Bajetta E, et al. Comparison of fulvestrant versus tamoxifen for the treatment of advanced breast cancer in postmenopausal women previously untreated with endocrine therapy: A multinational, double-blind, randomized trial. Journal of Clinical Oncology [Internet]. 2004 May;22(9):1605–13. Available from: https://doi.org/10.1200%2Fjco.2004.02.112

115. Clarke R, Tyson JJ, Dixon JM. Endocrine resistance in breast cancer – An overview and update. Molecular and Cellular Endocrinology [Internet]. 2015 Dec 15 [cited 2018 Apr 6];418:220–34. Available from: http://www.sciencedirect.com/science/article/pii/S0303720715301015

116. Johnston SR, Dowsett M. Aromatase inhibitors for breast cancer: Lessons from the laboratory. Nature Reviews Cancer [Internet]. 2003 Nov;3(11):821–31. Available from: https://doi.org/10.1038%2Fnrc1211

117. Kao YC, Cam LL, Laughton CA, Zhou D, Chen S. Binding characteristics of seven inhibitors of human aromatase: a site-directed mutagenesis study. Cancer research [Internet]. 1996 Aug;56(15):3451–60. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8758911

118. Results of the ATAC (arimidex, tamoxifen, alone or in combination) trial after completion of 5 years adjuvant treatment for breast cancer. The Lancet [Internet]. 2005 Jan;365(9453):60–2. Available from: https://doi.org/10.1016%2Fs0140-6736%2804%2917666-6

119. Garreau JR, Delamelena T, Walts D, Karamlou K, Johnson N. Side effects of aromatase inhibitors versus tamoxifen: The patients' perspective. Am J Surg. 2006 Oct;192(4):496–8.

120. Robertson JFR, Llombart-Cussac A, Rolski J, Feltl D, Dewar J, Macpherson E, et al. Activity of fulvestrant 500 mg versus anastrozole 1 mg as first-line treatment for advanced breast cancer: Results from the FIRST study. J Clin Oncol. 2009 Sep 20;27(27):4530–5.

121. Ellis MJ, Llombart-Cussac A, Feltl D, Dewar JA, Jasiówka M, Hewson N, et al. Fulvestrant 500 mg Versus Anastrozole 1 mg for the First-Line Treatment of Advanced Breast Cancer: Overall Survival Analysis From the Phase II FIRST Study. J Clin Oncol. 2015 Nov 10;33(32):3781–7.

122. Colleoni M, Montagna E. Neoadjuvant therapy for ER-positive breast cancers. Ann Oncol [Internet]. 2012 Sep 1 [cited 2018 Apr 6];23(suppl_10):x243–8. Available from: https://academic.oup.com/annonc/article/23/suppl_10/x243/206441

123. Ellis MJ, Tao Y, Luo J, A'Hern R, Evans DB, Bhatnagar AS, et al. Outcome Prediction for Estrogen Receptor–Positive Breast Cancer Based on Postneoadjuvant Endocrine Therapy Tumor Characteristics. J Natl Cancer

Inst [Internet]. 2008 Oct 1 [cited 2018 Apr 6];100(19):1380–8. Available from: https://academic.oup.com/jnci/article/100/19/1380/948060

124. Carlson RW, Henderson IC. Sequential hormonal therapy for metastatic breast cancer after adjuvant tamoxifen or anastrozole. Breast Cancer Res Treat. 2003;80 Suppl 1:S19–26; discussion S27–28.

125. Garcia T, Lehrer S, Bloomer WD, Schachter B. A Variant Estrogen Receptor Messenger Ribonucleic Acid is Associated with Reduced Levels of Estrogen Binding in Human Mammary Tumors. Mol Endocrinol [Internet]. 1988 Sep 1 [cited 2018 Apr 6];2(9):785–91. Available from: https://academic.oup.com/mend/article/2/9/785/2713641

126. Barone I, Brusco L, Fuqua SAW. Estrogen Receptor Mutations and Changes in Downstream Gene Expression and Signaling. Clin Cancer Res [Internet]. 2010 May 15 [cited 2018 Apr 6];16(10):2702–8. Available from: http://clincancerres.aacrjournals.org/content/16/10/2702

127. Murphy L, Cherlet T, Lewis A, Banu Y, Watson P. New insights into estrogen receptor function in human breast cancer. Annals of Medicine [Internet]. 2003 Jan 1 [cited 2018 Apr 6];35(8):614–31. Available from: https://doi.org/10.1080/07853890310014579

128. Merenbakh-Lamin K, Ben-Baruch N, Yeheskel A, Dvir A, Soussan-Gutman L, Jeselsohn R, et al. D538G Mutation in Estrogen Receptor-α: A Novel Mechanism for Acquired Endocrine Resistance in Breast Cancer. Cancer Res [Internet]. 2013 Dec 1 [cited 2018 Apr 6];73(23):6856–64. Available from: http://cancerres.aacrjournals.org/content/73/23/6856

129. Toy W, Shen Y, Won H, Green B, Sakr RA, Will M, et al. *ESR1* ligand-binding domain mutations in hormone-resistant breast cancer. Nature Genetics [Internet]. 2013 Dec [cited 2018 Apr 6];45(12):1439–45. Available from: https://www.nature.com/articles/ng.2822

130. Robinson DR, Wu Y-M, Vats P, Su F, Lonigro RJ, Cao X, et al. Activating *ESR1* mutations in hormone-resistant metastatic breast cancer. Nature Genetics [Internet]. 2013 Dec [cited 2018 Apr 6];45(12):1446–51. Available from: https://www.nature.com/articles/ng.2823

131. Jeselsohn R, Yelensky R, Buchwalter G, Frampton G, Meric-Bernstam F, Gonzalez-Angulo AM, et al. Emergence of Constitutively Active Estrogen Receptor-α Mutations in Pretreated Advanced Estrogen Receptor–Positive Breast Cancer. Clin Cancer Res [Internet]. 2014 Apr 1 [cited 2018 Apr 6];20(7):1757–67. Available from: http://clincancerres.aacrjournals.org/content/20/7/1757

132. Fuqua SAW, Gu G, Rechoum Y. Estrogen receptor (ER) α mutations in breast cancer: Hidden in plain sight. Breast Cancer Res Treat [Internet]. 2014 Feb 1 [cited 2018 Apr 6];144(1):11–9. Available from: https://link.springer.com/article/10.1007/s10549-014-2847-4

133. Oh A, List H-J, Reiter R, Mani A, Zhang Y, Gehan E, et al. The Nuclear Receptor Coactivator AIB1 Mediates Insulin-like Growth Factor I-induced Phenotypic Changes in Human Breast Cancer Cells. Cancer Res [Internet]. 2004 Nov 15 [cited 2018 Apr 6];64(22):8299–308. Available from: http://cancerres.aacrjournals.org/content/64/22/8299

134. Clarke R, Brünner N, Katz D, Glanz P, Dickson RB, Lippman ME, et al. The Effects of a Constitutive Expression of Transforming Growth Factor-α on the Growth of MCF-7 Human Breast Cancer Cells in Vitro and in Vivo. Mol Endocrinol [Internet]. 1989 Feb 1 [cited 2018 Apr 6];3(2):372–80. Available from: https:

//academic.oup.com/mend/article/3/2/372/2713880

135. Ojo D, Wei F, Liu Y, Wang E, Zhang H, Lin X, et al. Factors Promoting Tamoxifen Resistance in Breast Cancer via Stimulating Breast Cancer Stem Cell Expansion. Curr Med Chem. 2015;22(19):2360–74.

136. Johnston S, Pippen J, Pivot X, Lichinitser M, Sadeghi S, Dieras V, et al. Lapatinib Combined With Letrozole Versus Letrozole and Placebo As First-Line Therapy for Postmenopausal Hormone Receptor–Positive Metastatic Breast Cancer. JCO [Internet]. 2009 Nov 20 [cited 2018 Apr 6];27(33):5538–46. Available from: http://ascopubs.org/doi/10.1200/JCO.2009.23.3734

137. Kaufman B, Mackey JR, Clemens MR, Bapsy PP, Vaid A, Wardley A, et al. Trastuzumab Plus Anastrozole Versus Anastrozole Alone for the Treatment of Postmenopausal Women With Human Epidermal Growth Factor Receptor 2–Positive, Hormone Receptor–Positive Metastatic Breast Cancer: Results From the Randomized Phase III TAnDEM Study. JCO [Internet]. 2009 Nov 20 [cited 2018 Apr 6];27(33):5529–37. Available from: http://ascopubs.org/doi/10.1200/JCO.2008.20.6847

138. Burstein HJ, Cirrincione CT, Barry WT, Chew HK, Tolaney SM, Lake DE, et al. Endocrine Therapy With or Without Inhibition of Epidermal Growth Factor Receptor and Human Epidermal Growth Factor Receptor 2: A Randomized, Double-Blind, Placebo-Controlled Phase III Trial of Fulvestrant With or Without Lapatinib for Postmenopausal Women With Hormone Receptor–Positive Advanced Breast Cancer—CALGB 40302 (Alliance). JCO [Internet]. 2014 Dec 10 [cited 2018 Apr 6];32(35):3959–66. Available from: http://ascopubs.org/doi/10.1200/JCO.2014.56.7941

139. López□Knowles Elena, O'Toole Sandra A., McNeil Catriona M., Millar Ewan K.A., Qiu Min R., Crea Paul, et al. PI3K pathway activation in breast cancer is associated with the basal□like phenotype and cancer□specific mortality. International Journal of Cancer [Internet]. 2009 Aug 14 [cited 2018 Apr 6];126(5):1121–31. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/ijc.24831

140. Crowder RJ, Phommaly C, Tao Y, Hoog J, Luo J, Perou CM, et al. PIK3CA and PIK3CB Inhibition Produce Synthetic Lethality when Combined with Estrogen Deprivation in Estrogen Receptor–Positive Breast Cancer. Cancer Res [Internet]. 2009 May 1 [cited 2018 Apr 6];69(9):3955–62. Available from: http://cancerres.aacrjournals.org/content/69/9/3955

141. Gil EMC. Targeting the PI3K/AKT/mTOR pathway in estrogen receptor-positive breast cancer. Cancer Treatment Reviews [Internet]. 2014 Aug;40(7):862–71. Available from: https://doi.org/10.1016%2Fj.ctrv.2014.03.004

142. Carracedo A, Ma L, Teruya-Feldstein J, Rojo F, Salmena L, Alimonti A, et al. Inhibition of mTORC1 leads to MAPK pathway activation through a PI3K-dependent feedback loop in human cancer. J Clin Invest [Internet]. 2008 Sep 2 [cited 2018 Apr 6];118(9):3065–74. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2518073/

143. O'Reilly KE, Rojo F, She Q-B, Solit D, Mills GB, Smith D, et al. mTOR Inhibition Induces Upstream Receptor Tyrosine Kinase Signaling and Activates Akt. Cancer Res [Internet]. 2006 Feb 1 [cited 2018 Apr 6];66(3):1500–8. Available from: http://cancerres.aacrjournals.org/content/66/3/1500

144. Chakrabarty A, Sánchez V, Kuba MG, Rinehart C, Arteaga CL. Feedback upregulation of HER3 (ErbB3) expression and activity attenuates antitumor effect of PI3K inhibitors. PNAS [Internet]. 2012 Feb 21 [cited

2018 Apr 6];109(8):2718–23. Available from: http://www.pnas.org/content/109/8/2718

145. Chandarlapaty S, Sawai A, Scaltriti M, Rodrik-Outmezguine V, Grbovic-Huezo O, Serra V, et al. AKT Inhibition Relieves Feedback Suppression of Receptor Tyrosine Kinase Expression and Activity. Cancer Cell [Internet]. 2011 Jan 18 [cited 2018 Apr 6];19(1):58–71. Available from: http://www.sciencedirect.com/science/article/pii/S1535610810004332

146. Johnson MD, Zuo H, Lee K-H, Trebley JP, Rae JM, Weatherman RV, et al. Pharmacological Characterization of 4-hydroxy-<Emphasis Type="Italic">N</Emphasis>-desmethyl Tamoxifen, a Novel Active Metabolite of Tamoxifen. Breast Cancer Res Treat [Internet]. 2004 May 1 [cited 2018 Apr 6];85(2):151–9. Available from: https://link.springer.com/article/10.1023/B:BREA.0000025406.31193.e8

147. Clarke R, Leonessa F, Welch JN, Skaar TC. Cellular and molecular pharmacology of antiestrogen action and resistance. Pharmacol Rev. 2001 Mar;53(1):25–71.

148. Musgrove EA, Sutherland RL. Biological determinants of endocrine resistance in breast cancer. Nature Reviews Cancer [Internet]. 2009 Sep;9(9):631–43. Available from: https://doi.org/10.1038%2Fnrc2713

149. Piggott L, Silva A, Robinson T, Santiago-Gómez A, Simões BM, Becker M, et al. Acquired resistance of ER-positive breast cancer to endocrine treatment confers an adaptive sensitivity to TRAIL through posttranslational downregulation of c-FLIP. Clinical Cancer Research [Internet]. 2018 Jan;24(10):2452–63. Available from: https://doi.org/10.1158%2F1078-0432.ccr-17-1381

150. Chia S, Gradishar W, Mauriac L, Bines J, Amant F, Federico M, et al. Double-blind, randomized placebo controlled trial of fulvestrant compared with exemestane after prior nonsteroidal aromatase inhibitor therapy in postmenopausal women with hormone receptorPositive, advanced breast cancer: Results from EFECT. Journal of Clinical Oncology [Internet]. 2008 Apr;26(10):1664–70. Available from: https://doi.org/10.1200%2Fjco.2007.13.5822

151. Cottu P, Bieche I, Assayag F, Botty RE, Chateau-Joubert S, Thuleau A, et al. Acquired resistance to endocrine treatments is associated with tumor-specific molecular changes in patient-derived luminal breast cancer xenografts. Clinical Cancer Research [Internet]. 2014 Jun;20(16):4314–25. Available from: https://doi.org/10.1158%2F1078-0432.ccr-13-3230

152. Augusto TV, Correia-da-Silva G, Rodrigues CMP, Teixeira N, Amaral C. Acquired resistance to aromatase inhibitors: Where we stand! Endocrine-Related Cancer [Internet]. 2018 Mar;25(5):R283–301. Available from: https://doi.org/10.1530%2Ferc-17-0425

153. Demicheli R, Ardoino I, Boracchi P, Coradini D, Agresti R, Ferraris C, et al. Recurrence and mortality according to estrogen receptor status for breast cancer patients undergoing conservative surgery. Ipsilateral breast tumour recurrence dynamics provides clues for tumour biology within the residual breast. BMC Cancer. 2010 Nov 30;10:656.

154. Hiscox S, Jiang WG, Obermeier K, Taylor K, Morgan L, Burmi R, et al. Tamoxifen resistance in MCF7 cells promotes EMT-like behaviour and involves modulation of b-catenin phosphorylation. International Journal of Cancer [Internet]. 2006 Jan;118(2):290–301. Available from: https://doi.org/10.1002%2Fijc.21355

155. Yuan J, Liu M, Yang L, Tu G, Zhu Q, Chen M, et al. Acquisition of epithelial-mesenchymal transition phenotype in the tamoxifen-resistant breast cancer cell: A new role for g protein-coupled estrogen receptor in mediating tamoxifen resistance through cancer-associated fibroblast-derived fibronectin and

b1-integrin signaling pathway in tumor cells. Breast Cancer Research [Internet]. 2015 May;17(1). Available from: https://doi.org/10.1186%2Fs13058-015-0579-y

156. Işeri ÖD, Kars MD, Arpaci F, Atalay C, Pak I, Gündüz U. Drug resistant MCF-7 cells exhibit epithelial-mesenchymal transition gene expression pattern. Biomedicine & Pharmacotherapy [Internet]. 2011 Feb;65(1):40–5. Available from: https://doi.org/10.1016%2Fj.biopha.2010.10.004

157. Saleh SA, Mulla FA, Luqmani YA. Estrogen receptor silencing induces epithelial to mesenchymal transition in human breast cancer cells. Avraham HK, editor. PLoS ONE [Internet]. 2011 Jun;6(6):e20610. Available from: https://doi.org/10.1371%2Fjournal.pone.0020610

158. O'Brien CS, Farnie G, Howell SJ, Clarke RB. Breast cancer stem cells and their role in resistance to endocrine therapy. Hormones and Cancer [Internet]. 2011 Jan;2(2):91–103. Available from: https://doi.org/10.1007%2Fs12672-011-0066-6

159. Self-renewal of CD133hi cells by IL6/notch3 signalling regulates endocrine resistance in metastatic breast cancer. Nature Communications [Internet]. 2016 Feb;7:10442. Available from: https://doi.org/10.1038%2Fncomms10442

160. Osborne CK, Schiff R. Mechanisms of endocrine resistance in breast cancer. Annual Review of Medicine [Internet]. 2011 Feb;62(1):233–47. Available from: https://doi.org/10.1146%2Fannurev-med-070909-182917

161. Baselga J, Campone M, Piccart M, Burris HA, Rugo HS, Sahmoud T, et al. Everolimus in postmenopausal hormone-receptorPositive advanced breast cancer. New England Journal of Medicine [Internet]. 2012 Feb;366(6):520–9. Available from: https://doi.org/10.1056%2Fnejmoa1109653

162. Piccart M, Hortobagyi GN, Campone M, Pritchard KI, Lebrun F, Ito Y, et al. Everolimus plus exemestane for hormone-receptor-positive, human epidermal growth factor receptor-2-negative advanced breast cancer: Overall survival results from BOLERO-2. Annals of Oncology [Internet]. 2014 Sep;25(12):2357–62. Available from: https://doi.org/10.1093%2Fannonc%2Fmdu456

163. Finn RS, Crown JP, Lang I, Boer K, Bondarenko IM, Kulyk SO, et al. The cyclin-dependent kinase 4/6 inhibitor palbociclib in combination with letrozole versus letrozole alone as first-line treatment of oestrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/TRIO-18): A randomised phase 2 study. The Lancet Oncology [Internet]. 2015 Jan;16(1):25–35. Available from: https://doi.org/10.1016%2Fs1470-2045%2814%2971159-3

164. Uhr JW, Pantel K. Controversies in clinical cancer dormancy. Proceedings of the National Academy of Sciences [Internet]. 2011 Jul;108(30):12396–400. Available from: https://doi.org/10.1073%2Fpnas.1106613108

165. Sosa MS, Bragado P, Aguirre-Ghiso JA. Mechanisms of disseminated cancer cell dormancy: An awakening field. Nature Reviews Cancer [Internet]. 2014 Aug;14(9):611–22. Available from: https://doi.org/10.1038%2Fnrc3793

166. Dittmer J. Mechanisms governing metastatic dormancy in breast cancer. Seminars in Cancer Biology [Internet]. 2017 Jun;44:72–82. Available from: https://doi.org/10.1016%2Fj.semcancer.2017.03.006

167. Jatoi I, Proschan MA. Randomized Trials of Breast-conserving Therapy Versus Mastectomy for Primary Breast Cancer: A Pooled Analysis of Updated Results. American Journal of Clinical Oncology [Internet]. 2005 Jun 1 [cited 2018 Mar 19];28(3):289–94. Available from: https://insights.ovid.com/pubmed?pmid=

15923803

168. Fisher B, Costantino J, Redmond C, Poisson R, Bowman D, Couture J, et al. A Randomized Clinical Trial Evaluating Tamoxifen in the Treatment of Patients with Node-Negative Breast Cancer Who Have Estrogen-Receptor–Positive Tumors. New England Journal of Medicine [Internet]. 1989 Feb 23 [cited 2018 Mar 19];320(8):479–84. Available from: http://dx.doi.org/10.1056/NEJM198902233200802

169. van la Parra RFD, Kuerer HM. Selective elimination of breast cancer surgery in exceptional responders: Historical perspective and current trials. Breast Cancer Res [Internet]. 2016 [cited 2018 Mar 19];18. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4782355/

170. Thomaes T, Thomis M, Onkelinx S, Coudyzer W, Cornelissen V, Vanhees L. Reliability and validity of the ultrasound technique to measure the rectus femoris muscle diameter in older CAD-patients. BMC Medical Imaging [Internet]. 2012 Apr;12(1). Available from: https://doi.org/10.1186%2F1471-2342-12-7

171. Eri LM, Thomassen H, Brennhovd B, Håheim LL. Accuracy and repeatability of prostate volume measurements by transrectal ultrasound. Prostate Cancer and Prostatic Diseases [Internet]. 2002 Dec;5(4):273–8. Available from: https://doi.org/10.1038%2Fsj.pcan.4500568

172. Ulbricht L, Neves EB, Ripka WL, Romaneli EFR. Comparison between body fat measurements obtained by portable ultrasound and caliper in young adults. In: 2012 annual international conference of the IEEE engineering in medicine and biology society [Internet]. IEEE; 2012. Available from: https://doi.org/10.1109%2Fembc.2012.6346337

173. Reich B, Zalut T, Weiner SG. An international evaluation of ultrasound vs. computed tomography in the diagnosis of appendicitis. International Journal of Emergency Medicine [Internet]. 2011 Oct;4(1). Available from: https://doi.org/10.1186%2F1865-1380-4-68

174. Cha MJ, Lee MW, Cha DI, Kim J-H, Rhim H, Cho YK, et al. Size discrepancy between sonographic and computed tomographic/magnetic resonance imaging measurement of hepatocellular carcinoma. Journal of Ultrasound in Medicine [Internet]. 2013 Oct;32(10):1703–9. Available from: https://doi.org/10.7863%2Fultra.32.10.1703

175. Dowsett M, Smith IE, Ebbs SR, Dixon JM, Skene A, Griffith C, et al. Short-term changes in Ki-67 during neoadjuvant treatment of primary breast cancer with anastrozole or tamoxifen alone or combined correlate with recurrence-free survival. Clin Cancer Res. 2005 Jan 15;11(2 Pt 2):951s–8s.

176. Ellis MJ, Coop A, Singh B, Tao Y, Llombart-Cussac A, Jänicke F, et al. Letrozole inhibits tumor proliferation more effectively than tamoxifen independent of HER1/2 expression status. Cancer Res. 2003 Oct 1;63(19):6523–31.

177. Ge W-k, Yang B, Zuo W-s, Zheng G, Dai Y-q, Han C, et al. Evaluation of hormone receptor, human epidermal growth factor receptor-2 and Ki-67 with core needle biopsy and neoadjuvant chemotherapy effects in breast cancer patients. Thorac Cancer [Internet]. 2015 Jan [cited 2018 Mar 28];6(1):64–9. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4448469/

178. Zhou X, Zhang J, Yun H, Shi R, Wang Y, Wang W, et al. Alterations of biomarker profiles after neoadjuvant chemotherapy in breast cancer: Tumor heterogeneity should be taken into consideration. Oncotarget [Internet]. 2015 Sep 7 [cited 2018 Mar 28];6(34):36894–902. Available from: http://www.oncotarget.com/index.php?journal=oncotarget&page=article&op=view&path[]=5050&pubmed-

linkout=1

179. Greer LT, Rosman M, Mylander WC, Hooke J, Kovatich A, Sawyer K, et al. Does Breast Tumor Heterogeneity Necessitate Further Immunohistochemical Staining on Surgical Specimens? Journal of the American College of Surgeons [Internet]. 2013 Feb 1 [cited 2018 Mar 28];216(2):239–51. Available from: http://www.sciencedirect.com/science/article/pii/S1072751512012008

180. Chen X, Yuan Y, Gu Z, Shen K. Accuracy of estrogen receptor, progesterone receptor, and HER2 status between core needle and open excision biopsy in breast cancer: A meta-analysis. Breast Cancer Res Treat. 2012 Aug;134(3):957–67.

181. Li S, Yang X, Zhang Y, Fan L, Zhang F, Chen L, et al. Assessment accuracy of core needle biopsy for hormone receptors in breast cancer: A meta-analysis. Breast Cancer Research and Treatment [Internet]. 2012 Apr;135(2):325–34. Available from: https://doi.org/10.1007%2Fs10549-012-2063-z

182. Kim HS, Park S, Koo JS, Kim S, Kim JY, Nam S, et al. Risk Factors Associated with Discordant Ki-67 Levels between Preoperative Biopsy and Postoperative Surgical Specimens in Breast Cancers. PLOS ONE [Internet]. 2016 Mar 8 [cited 2018 Mar 28];11(3):e0151054. Available from: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0151054

183. Knutsvik G, Stefansson IM, Aziz S, Arnes J, Eide J, Collett K, et al. Evaluation of Ki67 Expression across Distinct Categories of Breast Cancer Specimens: A Population-Based Study of Matched Surgical Specimens, Core Needle Biopsies and Tissue Microarrays. PLOS ONE [Internet]. 2014 Nov 6 [cited 2018 Mar 28];9(11):e112121. Available from: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0112121

184. Chen X, Sun L, Mao Y, Zhu S, Wu J, Huang O, et al. Preoperative core needle biopsy is accurate in determining molecular subtypes in invasive breast cancer. BMC Cancer [Internet]. 2013 Aug 19 [cited 2018 Mar 28];13:390. Available from: https://doi.org/10.1186/1471-2407-13-390

185. Cruz LMDL, Harhay MO, Zhang P, Ugras S. Impact of neoadjuvant chemotherapy on breast cancer subtype: Does subtype change and, if so, how? Annals of Surgical Oncology [Internet]. 2018 Jul;25(12):3535–40. Available from: https://doi.org/10.1245%2Fs10434-018-6608-1

186. McAnena PF, McGuire A, Ramli A, Curran C, Malone C, McLaughlin R, et al. Breast cancer subtype discordance: Impact on post-recurrence survival and potential treatment options. BMC Cancer [Internet]. 2018 Feb;18(1). Available from: https://doi.org/10.1186%2Fs12885-018-4101-7

187. Jin X, Jiang Y-Z, Chen S, Yu K-D, Shao Z-M, Di G-H. Prognostic value of receptor conversion after neoadjuvant chemotherapy in breast cancer patients: A prospective observational study. Oncotarget [Internet]. 2015 Mar;6(11). Available from: https://doi.org/10.18632%2Foncotarget.3292

188. Montagna E, Bagnardi V, Viale G, Rotmensz N, Sporchia A, Cancello G, et al. Changes in PgR and ki-67 in residual tumour and outcome of breast cancer patients treated with neoadjuvant chemotherapy. Annals of Oncology [Internet]. 2014 Nov;26(2):307–13. Available from: https://doi.org/10.1093%2Fannonc%2Fmdu528

189. Dowsett M, Ebbs SR, Dixon JM, Skene A, Griffith C, Boeddinghaus I, et al. Biomarker changes during neoadjuvant anastrozole, tamoxifen, or the combination: Influence of hormonal status and HER-2 in breast cancerA study from the IMPACT trialists. Journal of Clinical Oncology [Internet]. 2005

Apr;23(11):2477–92. Available from: https://doi.org/10.1200%2Fjco.2005.07.559

190. Dowsett M, Smith IE, Ebbs SR, Dixon JM, Skene A, AHern R, et al. Prognostic value of ki67 expression after short-term presurgical endocrine therapy for primary breast cancer. JNCI: Journal of the National Cancer Institute. 2007 Jan;99(2):167–70.

191. Hofer SO, Shrayer D, Reichner JS, Hoekstra HJ, Wanebo HJ. Wound-induced tumor progression: A probable role in recurrence after tumor resection. Arch Surg. 1998 Apr;133(4):383–9.

192. Baker DG, Masterson TM, Pace R, Constable WC, Wanebo H. The influence of the surgical wound on local tumor recurrence. Surgery. 1989 Sep;106(3):525–32.

193. Hansbrough JF, Bender EM, Zapata-Sirvent R, Anderson J. Altered helper and suppressor lymphocyte populations in surgical patients. A measure of postoperative immunosuppression. Am J Surg. 1984 Sep;148(3):303–7.

194. Fisher B, Gunduz N, Coyle J, Rudock C, Saffer E. Presence of a growth-stimulating factor in serum following primary tumor removal in mice. Cancer Res. 1989 Apr 15;49(8):1996–2001.

195. Tagliabue E, Agresti R, Carcangiu ML, Ghirelli C, Morelli D, Campiglio M, et al. Role of HER2 in wound-induced breast carcinoma proliferation. Lancet. 2003 Aug 16;362(9383):527–33.

196. Chen X, Zhu S, Fei X, Garfield DH, Wu J, Huang O, et al. Surgery time interval and molecular subtype may influence Ki67 change after core needle biopsy in breast cancer patients. BMC Cancer [Internet]. 2015 Oct 30 [cited 2018 Mar 28];15. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4627413/

197. Jeselsohn RM, Werner L, Regan MM, Fatima A, Gilmore L, Collins LC, et al. Digital Quantification of Gene Expression in Sequential Breast Cancer Biopsies Reveals Activation of an Immune Response. Dahlman-Wright K, editor. PLoS ONE [Internet]. 2013 May;8(5):e64225. Available from: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0064225

198. Lee AHS, Happerfield LC, Bobrow LG, Millis RR. Angiogenesis and inflammation in ductal carcinoma in situ of the breast. The Journal of Pathology [Internet]. 1997 Feb;181(2):200–6. Available from: https://doi.org/10.1002%2F%28sici%291096-9896%28199702%29181%3A2%3C200%3A%3Aaid-path726%3E3.0.co%3B2-k

199. Mukhtar RA, Nseyo O, Campbell MJ, Esserman LJ. Tumor-associated macrophages in breast cancer as potential biomarkers for new treatments and diagnostics. Expert Review of Molecular Diagnostics [Internet]. 2011 Jan;11(1):91–100. Available from: https://doi.org/10.1586%2Ferm.10.97

200. Ripoli FL, Mohr A, Hammer SC, Willenbrock S, Hewicker-Trautwein M, Hennecke S, et al. A comparison of fresh frozen vs. formalin-fixed, paraffin-embedded specimens of canine mammary tumors via branched-DNA assay. International Journal of Molecular Sciences [Internet]. 2016 May;17(5):724. Available from: https://doi.org/10.3390%2Fijms17050724

201. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics (Oxford, England) [Internet]. 2007 Jan;8(1):118–27. Available from: http://biostatistics.oxfordjournals.org/content/8/1/118.full

202. Khatri P, Sirota M, Butte AJ. Ten years of pathway analysis: Current approaches and outstanding challenges. Ouzounis CA, editor. PLoS Computational Biology [Internet]. 2012 Feb;8(2):e1002375. Available

from: https://doi.org/10.1371%2Fjournal.pcbi.1002375

203. Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part i: Basic concepts and first analyses. British Journal of Cancer [Internet]. 2003 Jul;89(2):232–8. Available from: https://doi.org/10.1038%2Fsj.bjc.6601118

204. Kaplan EL, Meier P. Nonparametric Estimation from Incomplete Observations. Journal of the American Statistical Association [Internet]. 1958 Jun 1 [cited 2017 Aug 3];53(282):457–81. Available from: http://www.tandfonline.com/doi/abs/10.1080/01621459.1958.10501452

205. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. analysis and examples. British journal of cancer [Internet]. 1977 Jan;35(1):1–39. Available from: http://www.ncbi.nlm.nih.gov/pubmed/831755 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2025310

206. Loi S, Haibe-Kains B, Desmedt C, Wirapati P, Lallemand F, Tutt AM, et al. Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. BMC genomics [Internet]. 2008 May;9:239. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18498629 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2423197

207. Loi S, Haibe-Kains B, Desmedt C, Lallemand F, Tutt AM, Gillet C, et al. Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2007 Apr;25(10):1239–46.

208. Zhang Y, Sieuwerts AM, McGreevy M, Casey G, Cufer T, Paradiso A, et al. The 76-gene signature defines high-risk patients that benefit from adjuvant tamoxifen therapy. Breast cancer research and treatment. 2009 Jul;116(2):303–9.

209. Symmans WF, Hatzis C, Sotiriou C, Andre F, Peintinger F, Regitnig P, et al. Genomic index of sensitivity to endocrine therapy for breast cancer. J Clin Oncol. 2010 Sep 20;28(27):4111–9.

210. Smith IE, Johnson L, Dowsett M, Robertson JFR, Robison LE, Kokan JS, et al. Trial of perioperative endocrine therapy: Individualizing care (POETIC). Journal of Clinical Oncology [Internet]. 2011 May;29(15_suppl):TPS117–7. Available from: https://doi.org/10.1200%2Fjco.2011.29.15_suppl.tps117

211. Pennisi A, Kieber-Emmons T, Makhoul I, Hutchins L. Relevance of pathological complete response after neoadjuvant therapy for breast cancer. Breast Cancer: Basic and Clinical Research [Internet]. 2016 Jan;10:BCBCR.S33163. Available from: https://doi.org/10.4137%2Fbcbcr.s33163

212. Dowsett M, Nielsen TO, Hern RA, Bartlett J, Coombes RC, Cuzick J, et al. Assessment of ki67 in breast cancer: Recommendations from the international ki67 in breast cancer working group. JNCI Journal of the National Cancer Institute [Internet]. 2011 Sep;103(22):1656–64. Available from: https://doi.org/10.1093%2Fjnci%2Fdjr393

213. Guix M, Matos Granja N de, Meszoely I, Adkins TB, Wieman BM, Frierson KE, et al. Short preoperative treatment with erlotinib inhibits tumor cell proliferation in hormone receptorPositive breast cancers. Journal of Clinical Oncology [Internet]. 2008 Feb;26(6):897–906. Available from: https://doi.org/10.1200%2Fjco.2007.13.5939

214. Robertson JF, Nicholson RI, Bundred NJ, Anderson E, Rayter Z, Dowsett M, et al. Comparison of the short-term biological effects of 7α-[9-(4,4,5,5,5-pentafluoropentylsulfinyl)-nonyl]estra-1,3,5, (10)-triene-3,17β-

diol (faslodex) versus tamoxifen in postmenopausal women with primary breast cancer. Cancer Research [Internet]. 2001;61(18):6739–46. Available from: http://cancerres.aacrjournals.org/content/61/18/6739

215. Dowsett M, Bundred NJ, Decensi A, Sainsbury RC, Lu Y, Hills MJ, et al. Effect of raloxifene on breast cancer cell ki67 and apoptosis. Cancer Epidemiology and Prevention Biomarkers [Internet]. 2001;10(9):961–6. Available from: http://cebp.aacrjournals.org/content/10/9/961

216. Kitchen RR, Sabine VS, Sims AH, Macaskill EJ, Renshaw L, Thomas JS, et al. Correcting for intra-experiment variation in Illumina BeadChip data is necessary to generate robust gene-expression profiles. BMC genomics [Internet]. 2010 Jan;11(1):134. Available from: http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-11-134

217. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proceedings of the National Academy of Sciences of the United States of America [Internet]. 2003 Jul;100(14):8418–23. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=166244&tool=pmcentrez&rendertype=abstract

218. Desmedt C, Haibe-Kains B, Wirapati P, Buyse M, Larsimont D, Bontempi G, et al. Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes. Clinical cancer research : an official journal of the American Association for Cancer Research. 2008 Aug;14(16):5158–65.

219. Budczies J, Bockmayr M, Denkert C, Klauschen F, Lennerz JK, Györffy B, et al. Classical pathology and mutational load of breast cancer - integration of two worlds. The Journal of Pathology: Clinical Research [Internet]. 2015 Jul;1(4):225–38. Available from: https://doi.org/10.1002%2Fcjp2.25

220. Bjarnadottir O, Kimbung S, Johansson I, Veerla S, Jonsson M, Bendahl P-O, et al. Global transcriptional changes following statin treatment in breast cancer. Clinical Cancer Research [Internet]. 2015 Apr;21(15):3402–11. Available from: https://doi.org/10.1158%2F1078-0432.ccr-14-1403

221. Folgueira K. Gene expression profile of residual breast cancer after doxorubicin and cyclophosphamide neoadjuvant chemotherapy. Oncology Reports [Internet]. 2009 Aug;22(4). Available from: https://doi.org/10.3892%2For_00000503

222. Morrogh M, Andrade VP, Patil AJ, Qin L-X, Mo Q, Sakr R, et al. Differentially Expressed Genes in Window Trials are Influenced by the Wound-Healing Process: Lessons Learned from a Pilot Study with Anastrozole. Journal of Surgical Research [Internet]. 2012 Jul;176(1):121–32. Available from: http://www.sciencedirect.com/science/article/pii/S0022480411005166

223. Kelman Z. PCNA: Structure, functions and interactions. Oncogene [Internet]. 1997;14(6):629–40. Available from: https://doi.org/10.1038%2Fsj.onc.1200886

224. Madine MA, Swietlik M, Pelizon C, Romanowski P, Mills AD, Laskey RA. The roles of the MCM, ORC, and cdc6 proteins in determining the replication competence of chromatin in quiescent cells. Journal of Structural Biology [Internet]. 2000 Apr;129(2-3):198–210. Available from: https://doi.org/10.1006%2Fjsbi.2000.4218

225. Morgan DO. Principles of CDK regulation. Nature [Internet]. 1995 Mar;374(6518):131–4. Available from: https://doi.org/10.1038%2F374131a0

226. Cao H, Yang Z-X, Jiang G-Q. Expression and Clinical Significance of Activating Transcription Factor 3 in

Human Breast Cancer. Iranian Journal of Basic Medical Sciences. 2013 Nov;16(11):1151–4.

227. Morrogh M, Andrade VP, Patil AJ, Qin L-X, Mo Q, Sakr R, et al. Differentially expressed genes in window trials are influenced by the wound-healing process: Lessons learned from a pilot study with anastrozole. Journal of Surgical Research [Internet]. 2012 Jul;176(1):121–32. Available from: https://doi.org/10.1016%2Fj.jss.2011.05.058

228. Huang J, Qi R, Quackenbush J, Dauway E, Lazaridis E, Yeatman T. Effects of ischemia on gene expression. The Journal of surgical research. 2001 Aug;99(2):222–7.

229. Dash A, Maine IP, Varambally S, Shen R, Chinnaiyan AM, Rubin MA. Changes in Differential Gene Expression because of Warm Ischemia Time of Radical Prostatectomy Specimens. The American Journal of Pathology [Internet]. 2002 Nov;161(5):1743–8. Available from: http://www.sciencedirect.com/science/article/pii/S0002944010644513

230. Borgan E, Navon R, Vollan HKM, Schlichting E, Sauer T, Yakhini Z, et al. Ischemia caused by time to freezing induces systematic microRNA and mRNA responses in cancer tissue. Molecular oncology. 2011 Dec;5(6):564–76.

231. Aktas B, Sun H, Yao H, Shi W, Hubbard R, Zhang Y, et al. Global gene expression changes induced by prolonged cold ischemic stress and preservation method of breast cancer tissue. Molecular oncology [Internet]. 2014 May;8(3):717–27. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4048748&tool=pmcentrez&rendertype=abstract

232. López-Knowles E, Gao Q, Cheang MCU, Morden J, Parker J, Martin L-A, et al. Heterogeneity in global gene expression profiles between biopsy specimens taken peri-surgically from primary ER-positive breast carcinomas. Breast cancer research : BCR [Internet]. 2016 Jan;18(1):39. Available from: http://breast-cancer-research.biomedcentral.com/articles/10.1186/s13058-016-0696-2

233. Xie L, Li X, Wang Q, Zhou J, Shen J, Luo L, et al. Effects of core needle biopsy and subsequent neoadjuvant chemotherapy on molecular alterations and outcome in breast cancer. OncoTargets and Therapy [Internet]. 2018 Feb;Volume 11:677–85. Available from: https://doi.org/10.2147%2Fott.s145715

234. Azambuja E de, Holmes AP, Piccart-Gebhart M, Holmes E, Cosimo SD, Swaby RF, et al. Lapatinib with trastuzumab for HER2-positive early breast cancer (NeoALTTO): Survival outcomes of a randomised, open-label, multicentre, phase 3 trial and their association with pathological complete response. The Lancet Oncology [Internet]. 2014 Sep;15(10):1137–46. Available from: https://doi.org/10.1016%2Fs1470-2045%2814%2970320-1

235. DeMichele A, Yee D, Berry DA, Albain KS, Benz CC, Boughey J, et al. The Neoadjuvant Model Is Still the Future for Drug Development in Breast Cancer. Clinical cancer research : an official journal of the American Association for Cancer Research. 2015 Jul;21(13):2911–5.

236. Ong KR, Sims AH, Harvie M, Chapman M, Dunn WB, Broadhurst D, et al. Biomarkers of dietary energy restriction in women at increased risk of breast cancer. Cancer Prev Res (Phila Pa). 2009;2(8):720–31.

237. Wilson CL, Sims AH, Howell A, Miller CJ, Clarke RB. Effects of oestrogen on gene expression in epithelium and stroma of normal human breast tissue. Endocr Relat Cancer. 2006;13(2):617–28.

238. Eiermann W, Paepke S, Appfelstaedt J, Llombart-Cussac A, Eremin J, Vinholes J, et al. Preoperative treatment of postmenopausal breast cancer patients with letrozole: A randomized double-blind multicenter

study. Annals of Oncology [Internet]. 2001 Nov;12(11):1527–32. Available from: https://doi.org/10.1023%2Fa%3A1013128213451

239. Ellis MJ, Coop A, Singh B, Mauriac L, Llombert-Cussac A, Jänicke F, et al. Letrozole is more effective neoadjuvant endocrine therapy than tamoxifen for ErbB-1 and/or ErbB-2Positive, estrogen receptorPositive primary breast cancer: Evidence from a phase III randomized trial. Journal of Clinical Oncology [Internet]. 2001 Sep;19(18):3808–16. Available from: https://doi.org/10.1200%2Fjco.2001.19.18.3808

240. Fisher B, Jeong J-H, Dignam J, Anderson S, Mamounas E, Wickerham DL, et al. Findings from recent national surgical adjuvant breast and bowel project adjuvant studies in stage i breast cancer. JNCI Monographs [Internet]. 2001 Dec;2001(30):62–6. Available from: https://doi.org/10.1093%2Foxfordjournals.jncimonographs.a003463

241. Wallgren A, Bonetti M, Gelber R, Goldhirsch A, Castiglione-Gertsch M, Holmberg S, et al. Risk factors for locoregional recurrence among breast cancer patients: Results from international breast cancer study group trials i through VII. Journal of Clinical Oncology [Internet]. 2003 Apr;21(7):1205–13. Available from: https://doi.org/10.1200%2Fjco.2003.03.130

242. Demicheli R, Terenziani M, Bonadonna G. Estimate of tumor growth time for breast cancer local recurrences: Rapid growth after wake-up? Breast Cancer Research and Treatment [Internet]. 1998 Sep;51(2):133–7. Available from: https://doi.org/10.1023%2Fa%3A1005887422022

243. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database hallmark gene set collection. Cell Systems [Internet]. 2015 Dec;1(6):417–25. Available from: https://doi.org/10.1016%2Fj.cels.2015.12.004

244. Theodorou V, Stark R, Menon S, Carroll JS. GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility. Genome Research [Internet]. 2012 Nov;23(1):12–22. Available from: https://doi.org/10.1101%2Fgr.139469.112

245. Magnani L, Ballantyne EB, Zhang X, Lupien M. PBX1 genomic pioneer function drives ERa signaling underlying progression in breast cancer. Clurman BE, editor. PLoS Genetics [Internet]. 2011 Nov;7(11):e1002368. Available from: https://doi.org/10.1371%2Fjournal.pgen.1002368

246. Caceres II de, Cortes-Sempere M, Moratilla C, Machado-Pinilla R, Rodriguez-Fanjul V, Manguán-García C, et al. IGFBP-3 hypermethylation-derived deficiency mediates cisplatin resistance in non-small-cell lung cancer. Oncogene [Internet]. 2009 Dec;29(11):1681–90. Available from: https://doi.org/10.1038%2Fonc.2009.454

247. Crea F, Giovannetti E, Cortesi F, Mey V, Nannizzi S, Ruiz MIG, et al. Epigenetic mechanisms of irinotecan sensitivity in colorectal cancer cell lines. Molecular Cancer Therapeutics [Internet]. 2009 Jun;8(7):1964–73. Available from: https://doi.org/10.1158%2F1535-7163.mct-09-0027

248. Voso MT, Santini V, Finelli C, Musto P, Pogliani E, Angelucci E, et al. Valproic acid at therapeutic plasma levels may increase 5-azacytidine efficacy in higher risk myelodysplastic syndromes. Clinical Cancer Research [Internet]. 2009 Jul;15(15):5002–7. Available from: https://doi.org/10.1158%2F1078-0432.ccr-09-0494

249. Ramalingam SS, Parise RA, Ramananthan RK, Lagattuta TF, Musguire LA, Stoller RG, et al. Phase i and pharmacokinetic study of vorinostat, a histone deacetylase inhibitor, in combination with car-

boplatin and paclitaxel for advanced solid malignancies. Clinical Cancer Research [Internet]. 2007 May;13(12):3605–10. Available from: https://doi.org/10.1158%2F1078-0432.ccr-07-0162

250. Rasheed W, Bishton M, Johnstone RW, Prince HM. Histone deacetylase inhibitors in lymphoma and solid malignancies. Expert Review of Anticancer Therapy [Internet]. 2008 Mar;8(3):413–32. Available from: https://doi.org/10.1586%2F14737140.8.3.413

251. Zhang Y, Ishida CT, Ishida W, Lo S-FL, Zhao J, Shu C, et al. Combined HDAC and bromodomain protein inhibition reprograms tumor cell metabolism and elicits synthetic lethality in glioblastoma. Clinical Cancer Research [Internet]. 2018 May; Available from: https://doi.org/10.1158%2F1078-0432.ccr-18-0260

252. Bolden JE, Shi W, Jankowski K, Kan C-Y, Cluse L, Martin BP, et al. HDAC inhibitors induce tumor-cell-selective pro-apoptotic transcriptional responses. Cell Death & Disease [Internet]. 2013 Feb;4(2):e519–9. Available from: https://doi.org/10.1038%2Fcddis.2013.9

253. Yardley DA, Ismail-Khan RR, Melichar B, Lichinitser M, Munster PN, Klein PM, et al. Randomized phase II, double-blind, placebo-controlled study of exemestane with or without entinostat in postmenopausal women with locally recurrent or metastatic estrogen receptor-positive breast cancer progressing on treatment with a nonsteroidal aromatase inhibitor. Journal of Clinical Oncology [Internet]. 2013 Jun;31(17):2128–35. Available from: https://doi.org/10.1200%2Fjco.2012.43.7251

254. Yeruva SLH, Zhao F, Miller KD, Tevaarwerk AJ, Wagner LI, Gray RJ, et al. E2112: Randomized phase iii trial of endocrine therapy plus entinostat/placebo in patients with hormone receptor-positive advanced breast cancer. npj Breast Cancer [Internet]. 2018 Jan;4(1). Available from: https://doi.org/10.1038%2Fs41523-017-0053-3

255. Magnani L, Stoeck A, Zhang X, Lanczky A, Mirabella AC, Wang T-L, et al. Genome-wide reprogramming of the chromatin landscape underlies endocrine therapy resistance in breast cancer. Proceedings of the National Academy of Sciences [Internet]. 2013 Apr;110(16):E1490–9. Available from: https://doi.org/10.1073%2Fpnas.1219992110

256. Stone A, Valdés-Mora F, Gee JMW, Farrow L, McClelland RA, Fiegl H, et al. Tamoxifen-induced epigenetic silencing of oestrogen-regulated genes in anti-hormone resistant breast cancer. Futscher BW, editor. PLoS ONE [Internet]. 2012 Jul;7(7):e40466. Available from: https://doi.org/10.1371%2Fjournal.pone.0040466

257. Mehta RJ, Jain RK, Leung S, Choo J, Nielsen T, Huntsman D, et al. FOXA1 is an independent prognostic marker for ER-positive breast cancer. Breast Cancer Research and Treatment [Internet]. 2011 Apr;131(3):881–90. Available from: https://doi.org/10.1007%2Fs10549-011-1482-6

258. Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature [Internet]. 2012 Jan;481(7381):389–93. Available from: https://doi.org/10.1038%2Fnature10730

259. Connolly R, Stearns V. Epigenetics as a therapeutic target in breast cancer. Journal of Mammary Gland Biology and Neoplasia [Internet]. 2012 Jul;17(3-4):191–204. Available from: https://doi.org/10.1007%2Fs10911-012-9263-3

260. Parrella P. Epigenetic signatures in breast cancer: Clinical perspective. Breast Care [Internet]. 2010;5(2):66–73. Available from: https://doi.org/10.1159%2F000309138

261. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: A new bio-informatics tool for biomarker assessment and

outcome-based cut-point optimization. Clin Cancer Res. 2004 Nov 1;10(21):7252–9.

262. Motakis E, Ivshina AV, Kuznetsov VA. Data-driven approach to predict survival of cancer patients. IEEE Engineering in Medicine and Biology Magazine. 2009 Jul;28(4):58–66.

263. Mizuno H, Kitada K, Nakai K, Sarai A. PrognoScan: A new database for meta-analysis of the prognostic value of genes. BMC Medical Genomics [Internet]. 2009 Apr 24 [cited 2017 Jul 17];2:18. Available from: http://dx.doi.org/10.1186/1755-8794-2-18

264. Cox DR. Regression Models and Life-Tables. Journal of the Royal Statistical Society Series B (Methodological) [Internet]. 1972;34(2):187–220. Available from: http://links.jstor.org/sici?sici=0035-9246%25281972%252934%253A2%253C187%253ARMAL%253E2.0.CO%253B2-6

265. Lazic SE. Ranking, selecting, and prioritising genes with desirability functions. PeerJ [Internet]. 2015 Nov;3:e1444. Available from: https://doi.org/10.7717%2Fpeerj.1444

266. Kapp AV, Jeffrey SS, Langerød A, Børresen-Dale A-L, Han W, Noh D-Y, et al. Discovery and validation of breast cancer subtypes. BMC genomics [Internet]. 2006 Jan;7(1):231. Available from: http://www.biomedcentral.com/1471-2164/7/231

267. Dean-Colomb W, Esteva FJ. Her2-positive breast cancer: Herceptin and beyond. European Journal of Cancer [Internet]. 2008 Dec 1 [cited 2017 Aug 3];44(18):2806–12. Available from: http://www.sciencedirect.com/science/article/pii/S0959804908007156

268. Miyoshi Y, Iwao K, Egawa C, Noguchi S. Association of centrosomal kinaseSTK15/BTAK mRNA expression with chromosomal instability in human breast cancers. International Journal of Cancer [Internet]. 2001;92(3):370–3. Available from: https://doi.org/10.1002%2Fijc.1200

269. Willems E, Dedobbeleer M, Digregorio M, Lombard A, Lumapat PN, Rogister B. The functional diversity of aurora kinases: A comprehensive review. Cell Division [Internet]. 2018 Sep;13(1). Available from: https://doi.org/10.1186%2Fs13008-018-0040-6

270. Long PM, Stradecki HM, Minturn JE, Wesley UV, Jaworski DM. Differential aminoacylase expression in neuroblastoma. International Journal of Cancer [Internet]. 2011 Apr;129(6):1322–30. Available from: https://doi.org/10.1002%2Fijc.25798

271. Agalliu I, Luciano MS, Mirelman A, Giladi N, Waro B, Aasly J, et al. Higher frequency of certain cancers inLRRK2G2019S mutation carriers with parkinson disease. JAMA Neurology [Internet]. 2015 Jan;72(1):58. Available from: https://doi.org/10.1001%2Fjamaneurol.2014.1973

272. Emma MR, Iovanna JL, Bachvarov D, Puleio R, Loria GR, Augello G, et al. NUPR1, a new target in liver cancer: Implication in controlling cell growth, migration, invasion and sorafenib resistance. Cell Death & Disease [Internet]. 2016 Jun;7(6):e2269–9. Available from: https://doi.org/10.1038%2Fcddis.2016.175

273. Zheng Z, Park JY, Guillemette C, Schantz SP, Lazarus P. Tobacco carcinogen-detoxifying enzyme UGT1A7 and its association with orolaryngeal cancer risk. JNCI Journal of the National Cancer Institute [Internet]. 2001 Sep;93(18):1411–8. Available from: https://doi.org/10.1093%2Fjnci%2F93.18.1411

274. Szász AM, Lánczky A, Nagy, Förster S, Hark K, Green JE, et al. Cross-validation of survival associated biomarkers in gastric cancer using transcriptomic data of 1,065 patients. Oncotarget [Internet]. 2016

Jun;7(31). Available from: https://doi.org/10.18632%2Foncotarget.10337

275. Budczies J, Klauschen F, Sinn BV, Győrffy B, Schmitt WD, Darb-Esfahani S, et al. Cutoff finder: A comprehensive and straightforward web application enabling rapid biomarker cutoff optimization. Diest P van, editor. PLoS ONE [Internet]. 2012 Dec;7(12):e51862. Available from: https://doi.org/10.1371%2Fjournal.pone.0051862

276. Yoon S, Eom GH. HDAC and HDAC inhibitor: From cancer to cardiovascular diseases. Chonnam Medical Journal [Internet]. 2016;52(1):1. Available from: https://doi.org/10.4068%2Fcmj.2016.52.1.1

277. Calapre L, Warburton L, Millward M, Ziman M, Gray ES. Circulating tumour DNA (ctDNA) as a liquid biopsy for melanoma. Cancer Letters [Internet]. 2017 Sep;404:62–9. Available from: https://doi.org/10.1016%2Fj.canlet.2017.06.030

278. Church TR, Wandell M, Lofton-Day C, Mongin SJ, Burger M, Payne SR, et al. Prospective evaluation of methylatedSEPT9in plasma for detection of asymptomatic colorectal cancer. Gut [Internet]. 2013 Feb;63(2):317–25. Available from: https://doi.org/10.1136%2Fgutjnl-2012-304149

279. Warton K, Samimi G. Methylation of cell-free circulating DNA in the diagnosis of cancer. Frontiers in Molecular Biosciences [Internet]. 2015 Apr;2. Available from: https://doi.org/10.3389%2Ffmolb.2015.00013

280. Wang Y, Waters J, Leung ML, Unruh A, Roh W, Shi X, et al. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. Nature [Internet]. 2014 Jul;512(7513):155–60. Available from: https://doi.org/10.1038%2Fnature13600

281. Creighton CJ, Massarweh S, Huang S, Tsimelzon A, Hilsenbeck SG, Osborne CK, et al. Development of resistance to targeted therapies transforms the clinically associated molecular profile subtype of breast tumor xenografts. Cancer Research [Internet]. 2008 Sep;68(18):7493–501. Available from: https://doi.org/10.1158%2F0008-5472.can-08-1404

282. Holen I, Speirs V, Morrissey B, Blyth K. In vivomodels in breast cancer research: Progress, challenges and future directions. Disease Models & Mechanisms [Internet]. 2017 Apr;10(4):359–71. Available from: https://doi.org/10.1242%2Fdmm.028274